

**UNIVERSIDAD COMPLUTENSE DE MADRID**

**FACULTAD DE FARMACIA**



**TESIS DOCTORAL**

**Estudio y caracterización del microbioma viral en pacientes  
pediátricos con infección aguda del tracto respiratorio**

MEMORIA PARA OPTAR AL GRADO DE DOCTORA

PRESENTADA POR

**María de la Montaña Iglesias Caballero**

Directores

**Inmaculada Casas Flecha  
Juan Francisco Pozo Sánchez**

**Madrid**

# **ESTUDIO Y CARACTERIZACIÓN DEL MICROBIOMA VIRAL EN PACIENTES PEDIÁTRICOS CON INFECCIÓN AGUDA DEL TRACTO RESPIRATORIO**

Tesis Doctoral

**María de la Montaña Iglesias Caballero**

**Madrid 2019**

Dirigida por:

**Inmaculada Casas Flecha**

**Juan Francisco Pozo Sánchez**

Facultad de Farmacia







Universidad Complutense de Madrid  
Departamento de Microbiología  
Facultad de Farmacia  
Programa de Doctorado: Microbiología

## **TESIS DOCTORAL**

**ESTUDIO Y CARACTERIZACIÓN DEL MICROBIOMA VIRAL EN PACIENTES PEDIÁTRICOS  
CON INFECCIÓN AGUDA DEL TRACTO RESPIRATORIO**

**Presentada por María de la Montaña Iglesias Caballero para optar al grado de Doctora**

**Dirigida por la Dra. Inmaculada Casas Flecha y el Dr. Juan Francisco Pozo Sánchez**

Madrid 2019







## DECLARACIÓN DE AUTORÍA Y ORIGINALIDAD DE LA TESIS PRESENTADA PARA OBTENER EL TÍTULO DE DOCTOR

Dña. María de la Montaña Iglesias Caballero, estudiante en el Programa de Doctorado de Microbiología Y Parasitología, de la Facultad de Farmacia de la Universidad Complutense de Madrid, como autor/a de la tesis presentada para la obtención del título de Doctor y titulada: "Estudio y caracterización del microbioma viral en pacientes pediátricos con infección aguda del tracto respiratorio" y dirigida por: Inmaculada Casas Flecha y Francisco Pozo Sánchez

### DECLARO QUE:

La tesis es una obra original que no infringe los derechos de propiedad intelectual ni los derechos de propiedad industrial u otros, de acuerdo con el ordenamiento jurídico vigente, en particular, la Ley de Propiedad Intelectual (R.D. legislativo 1/1996, de 12 de abril, por el que se aprueba el texto refundido de la Ley de Propiedad Intelectual, modificado por la Ley 2/2019, de 1 de marzo, regularizando, aclarando y armonizando las disposiciones legales vigentes sobre la materia), en particular, las disposiciones referidas al derecho de cita.

Del mismo modo, asumo frente a la Universidad cualquier responsabilidad que pudiera derivarse de la autoría o falta de originalidad del contenido de la tesis presentada de conformidad con el ordenamiento jurídico vigente.

En Madrid, a 19 de Septiembre de 2019

Fdo:



***“El viaje no termina jamás. Solo los viajeros terminan. Y también ellos pueden subsistir en memoria, en recuerdo, en narración... El objetivo de un viaje es sólo el inicio de otro viaje”.***

**José Saramago**



*A todos los que hicieron que llegara hasta aquí.*

*A todo lo que me hizo llegar hasta aquí.*

*Gracias de corazón.*



# Agradecimientos

Aunque no lo parezca este apartado es de lo más difíciles de esta tesis. La existencia de este tocho es posible gracias a tanta gente que no caben en estas páginas y es que si me llevo algo bueno de estos 4 años es toda la gente que me ha acompañado.

El primer párrafo debe ser para ti Inma, no sé si voy a saber expresar todo lo que te agradezco la paciencia y la ayuda. La experiencia de hacer una tesis teniendo a una mentora más que una jefa, que te acompañe una persona que no solo te enseña a nivel profesional si no que te aporta a nivel personal, ha sido probablemente lo mejor de estos 4 años. Me llevo tu ejemplo personal y profesional para el futuro (la persona con las mejores ideas de la virología patria), tu bolso chino (que no, que te lo devuelvo) y una cita dentro de 10 años para las risas (¿cueces o enriqueces?).

A ti Paco te agradezco tu atención y tu ayuda cada vez que asalto tu despacho. Gracias también por poner siempre otro punto de vista con el que aprender y enriquecerme como profesional y como persona. Siempre me voy a acordar de tus bromas y tu “pues quitando todo eso estamos estupendamente”.

A Moni (Moni Moni mono de Moni) y Mar (Nornolinero) os tengo que agradecer desde las enseñanzas más básicas para sobrevivir en el laboratorio hasta soportarme tésica perdía en el despachito. Gracias por los cuidados maternos, los consejos, las risas y siempre estar ahí como mi red de seguridad.

A Ana, Sara, Sheila y María agradeceros la ayuda, los ratitos de conversaciones sobre viajes, Harry Potter o el tema favorito de conversación “la comida” que hicieron más llevaderos los días.

Mil gracias a la Unidad de Genómica del ISCIII y en especial a ti Pilar, te agradezco que esta tesis tenga peso. Gracias por tu generosidad enseñándome los secretos de la secuenciación masiva, por hacer que cada vez que subía a genómica me sintiese como en casa, por traer el robot a mi vida y por todo el trabajo que gracias a ti ha salido adelante.

Mari Paz, artista de renombre en permanente gira mundial para hablar de los “argovirus”, que paciencia me has tenido en el verano de la agonía. Gracias por el cariño, los buenos consejos y sobre todo gracias por las conversaciones interesantes y las risas que me animaban mucho.

A ti Juan te agradezco mucho aquella noche murcielaguil que ha sido de las cosas más especiales de la tesis.

Unai te debo una dedicatoria en condiciones en los agradecimientos, unos 3000 litros de birra y un yate para que lo aparques en la ría. Gracias por ser mi Siri, por toda la ayuda, por ser en remoto un porcentaje altísimo de este trabajo y por todos los ánimos y apoyo que siempre me das. Agradecida por haber encontrado un amigo como tú.



Y no hubiese sido posible lograr esto sin la confianza de mi querido Maestro Jedi Lucas, gracias por la oportunidad que me diste, gracias por domesticar a este Manolito salvaje (sin daun) y gracias por los “ánimos y la confianza” pero sobre todo por las risas, los libros, las series y la música. Y al beatle de mollina, el del papel de orillo, a ti mi querido Ernesto decirte que me he acordado de ti muchas veces en este tiempo, que eres de las mejores personas que me he cruzado en el gremio y que ojalá te salga todo redondo porque eres el mejor.

Gracias a “No Pasa Nada” (Franin, Tani, Vio, Lauri...) por hacer que pasen cosas buenas y llenar las terrazas Vallekanas de diversión. Gracias a mi shashipeople (Aurori, Vero, Pedrete y María) por hacer de un año duro un año inolvidable. Y a mis biólogos granatensis (Sara, Alfredo, Maríalucena, Jorge, Yedra, Aran, Vito...) gracias por lo apuntes y las aventuras. Que voy a ser doctora!, si nos lo dicen un día en la puerta de ciencias saltándome Fisio Vegetal con las cariocas y la cerveza nos reímos. A mis granatensis de la revolución (Carlos, Lucía, Domi,...) gracias por los debates, la amistad y por hacerme mejor persona.

Muchas gracias papá y mamá por todas las cosas que me habéis enseñado silenciosamente todos estos años con vuestro ejemplo. Todos los valores, las oportunidades y los ánimos que me habéis dado desde que cantaba “Notoliti” hasta hoy es la gasolina que me mueve y me lleva a buen puerto. Nos quedan unos cuantos hitos vitales que superar y disfrutar juntos. A mi tata gracias por ser y estar, eres mi persona favorita del mundo (lo sabes, dame un besito). Gracias por tu humor (capasao), por echarle gasolina a mis tonterías y por tu apoyo. HOY TUMBO.

Y por último gracias ti Quique, el mejor compañero del mundo. Gracias por entender la locura que es la tesis y apoyarme a muerte siempre. Gracias por los cuidados, los ánimos y ser mi conductor a las 5 de la mañana. Haces que mi vida sea maravillosa y quiero que siga así 3300 tesis más.

A tod@s l@s que me acompañáis en este día gracias.

*María Iglesias-Caballero.*

*Laboratorio de Referencia de Bares y Circovirus de los afters.*

# CONTENIDO GENERAL

	Páginas
Resumen/Summary	3-7
Índice	9-14
Abreviaturas y unidades de medida	15-19
Apéndice de figuras y tablas	20-27
Introducción	29-52
Hipótesis y Objetivos	51-54
Materiales y métodos	55-106
Resultados y discusión	107-204
Epílogo	205-212
Conclusiones	213-216
Bibliografía	217-234
Congresos y artículos	235-273



## **Resumen (500-1000 palabras)**

El avance de la tecnología de secuenciación de nueva generación supone una oportunidad para la mejora de la investigación y el diagnóstico, abriendo un horizonte de retos y avances para un laboratorio de referencia.

La posibilidad de caracterizar con un único sistema la comunidad viral presente en una muestra biológica y detectar virus, conocidos o no, sin necesidad de cultivar, clonar o restringir la técnica a un virus concreto supone una gran ventaja.

El empleo de este sistema en el estudio del tracto respiratorio es un hilo conector que nos lleva desde la puesta a punto de la técnica hasta la comprensión de la ecología de la infección respiratoria y basándose en esto hemos desarrollado nuestros objetivos.

El estudio del viroma respiratorio de pacientes pediátricos nos permitió desarrollar los protocolos de enriquecimiento por tres métodos diferentes, de enriquecimiento mediante captura por sondas y la amplificación independiente de secuencia por oligonucleótido único. Estos métodos han sido posteriormente aplicados en el estudio de viroma respiratorio por una parte en niños y en personas adultas, y por otra en hospedadores diferentes al ser humano como son los murciélagos. Este sistema ha sido el método desarrollado y empleado fundamentalmente en este trabajo.

La puesta a punto de técnicas para el estudio de viroma respiratorio de pacientes pediátricos con infección respiratoria permitió la identificación de patógenos en muestras sin diagnosticar por métodos moleculares. Al incluir un grupo control sano se pudo determinar que los virus detectados en pacientes con infección respiratoria no circulaban en población sana. A su vez el empleo de este grupo permitió la caracterización de viroma respiratorio de pacientes pediátricos sanos.

Los resultados del estudio de rinovirus mostraron que los pacientes analizados estaban infectados en su mayoría por rinovirus A. En este análisis destacó la presencia de rinovirus junto con el virus circular Torque Teno en pacientes hospitalizados. La utilización de muestras de seguimiento de determinados pacientes permitió analizar la presencia del virus a lo largo del tiempo, determinar el período de excreción viral y los cambios genéticos que se producen a lo largo del tiempo en unos virus tan heterogéneos como son los rinovirus.

La caracterización de viroma de una paciente con una infección respiratoria grave permitió el descubrimiento de un nuevo virus circular. La utilización de la técnica de amplificación por círculo rodante permitió la obtención del genoma completo y la demostración de la presencia de las estructuras moleculares que definen genéticamente a los virus CRESS-DNA: el nomámero conservado y los motivos RCR y Walker A, B y C. La obtención de genoma completo ha permitido el diseño de sistemas de amplificación genómica aplicados al estudio y secuenciación de estos nuevos virus y gracias a ellos se identificaron virus circulares similares que podrían pertenecer a la recientemente publicada familia *Redondoviridae*. El estudio de caracterización de estos nuevos virus CRESS-DNA describió una baja presencia en muestras pediátricas y su detección en 5 pacientes con neumonía severa negativos a otros virus. En muestras pediátricas del grupo control negativo no fue detectado este virus.

La implementación de estas técnicas de estudio genómico no es importante únicamente a nivel de la investigación de viroma respiratorio, si no que su uso en el Laboratorio de Referencia Nacional de Gripe tiene un amplísimo potencial para la mejora del tipado molecular y la epidemiología de gripe. La puesta a punto de protocolos de secuenciación de los ocho segmentos que componen el genoma de los virus gripales permitió el estudio y caracterización de los virus de una paciente pediátrica fallecida sin otra causa que una infección gripal y su hermano, ambos coinfectados con virus de la gripe tipo A y tipo B. Inicialmente y por los métodos tradicionales no fue posible detectar la coinfección en la paciente fallecida, demostrando el potencial diagnóstico y la aplicación de este método en el estudio epidemiológico de la gripe.

El sistema desarrollado para el estudio de viroma respiratorio puede ser empleado directamente en el estudio de cualquier muestra respiratoria incluyendo muestras procedentes del ser humano y de animales. El estudio de la presencia de virus en murciélagos ibéricos resultó muy positiva con el descubrimiento de nuevos mastadenovirus, considerándose nuevas especies potenciales de adenovirus.

El descubrimiento de nuevos virus y en concreto de adenovirus directamente de muestra respiratoria, sin necesidad de cultivo, es todo un reto en la actualidad. La secuenciación del genoma completo de virus ADN de gran tamaño también lo es. La técnica de enriquecimiento mediante captura por sondas es una alternativa tecnológica que puede mejorar los resultados de los estudios genómicos en el caso de los virus ADN de gran tamaño. Un estudio comparativo entre los métodos de enriquecimiento de secuencia puede ser crucial a la hora de establecer el

mejor sistema de análisis de estos virus. Utilizando diferentes protocolos de enriquecimiento directamente en muestras biológicas se han obtenido fragmentos de alrededor de 17000 nucleótidos y en el momento de la escritura de esta tesis se sigue trabajando para la obtención de la secuencia completa de los nuevos mastadenovirus detectados en murciélagos ibéricos. La obtención de la secuencia genómica completa depende de factores ajenos a los sistemas de estudio como son: el estado de conservación de las muestras, que en este caso fueron tomadas en las campañas de captura de campo, o la baja carga viral en la muestra respiratoria tomadas de individuos aparentemente sanos.

El impacto de los métodos desarrollados en el estudio genómico de virus aplicados al diagnóstico de pacientes con infección respiratoria aguda y a la caracterización intrínseca de su viroma durante el desarrollo del cuadro clínico asociado, permite profundizar en la caracterización de los virus detectados como antes nunca se había realizado. El papel de los virus en el compartimiento respiratorio y sus relaciones ecológicas aportan una valiosísima información que de manera directa se traduce en un aumento de las posibilidades de estudio de virus en los Laboratorios de Referencia que cambiarán los paradigmas de la infección respiratoria tal y como los conocemos hoy.

## Summary (500-1000 palabras)

Next-generation sequencing development brings a new horizon of challenges and progress in a reference laboratory providing the opportunity to improve research and diagnostics.

The characterization of the viral community with one single method with the ability to find known viruses or not, without the need of cell culturing, cloning or biased molecular methods, is a great advantage. The use of NGS methods is a connecting thread that leads us from the evolvement of the protocols to a better knowledge of respiratory infection. Based on these points the objectives were developed.

The pediatric virome study allowed the standardization of the viral enrichment protocol by 3 steps, the virome capture sequencing platform and the sequence independent single primer amplification protocols. These methods were used in the respiratory virome studies of adults and bats.

The analysis of the pediatric respiratory virome identified viral pathogens in samples without diagnostic by molecular methods. The absence in the healthy donor samples of the viruses identified in patients with respiratory infection proved the role of these viruses in respiratory infection and characterized the viruses conforming the respiratory virome in health.

The role of rhinovirus in virome study showed more infections by rhinovirus A in our patients than with the other group of rhinoviruses. The analysis highlighted a possible association between rhinovirus and the circular virus Torque Teno virus in hospitalized patients. The use of follow-up samples allowed the characterization of the excretion period during rhinovirus infection.

A new circular virus was found in the virome of patients with a severe acute respiratory infection. The use of rolling circle amplification allowed the obtaining of full genome sequences where conserved CRESS-DNA viruses structures as the conserved nonanucleotide or RCR and Walker motifs were defined. The sequences were used for designing a study PCRs that allowed the detection of more similar viruses that could be part of the recently described family *Redondoviridae*. The study of the viral distribution in the population showed a low viral presence in pediatric patients. The presence of negative patients for other viruses with severe pneumonia positives for this circular virus was proved. These viruses were not detected in the control group.

The implementation of next-generation sequencing methods is not only relevant in respiratory virome research. The use of this technology in the National Influenza Reference Laboratory has a huge potential to improve the molecular typing and flu epidemiology. The implementation of protocols for influenza whole genome sequencing allowed characterization of the viruses detected in a deceased patient without any other cause known except flu infection. The patient and her brother were co-infected with influenza A and B viruses. Initially, using traditional methods for diagnosis the co-infection was not detected in the deceased patient proving the diagnostic potential of this method and its application in the epidemiologic study.

The workflow developed for respiratory virome study can be used for any type of sample including animals and human. The screening of the presence of viruses in Iberian bats revealed the presence of a high number of mastadenovirus that were potentially new species of adenovirus. The description of new viruses, and specifically big DNA virus as adenovirus, directly from respiratory samples without the need of cell culture is a big challenge. Viral enrichment with capture probes could improve the results in the case of these viruses.

A comparative study among the viral enrichment methods can be essential to establish the best analysis system for adenovirus. After the used of capture probes was obtained the biggest fragment at this moment (16884 nucleotides) and currently the method is still being improved to achieve the bat mastadenovirus full genome sequencing.

The impact of these methods for viral genomic studies in the diagnosis of patients with acute respiratory infection and the virome characterization during the clinical condition, allowed for a better understanding of the viruses as never before. The role of the viruses in the respiratory tract and the ecological relations among them, gave us important information that increase the possibilities of a better viral characterization in a Reference Laboratory. The information provided by this method can change the paradigms of the respiratory infection, as we know.





## Índice:

<b>Abreviaturas y unidades de medida.</b>	<b>15-19</b>
<b>Apéndice de figuras y tablas</b>	<b>20-27</b>
<b>1. INTRODUCCIÓN</b>	
1.1. Desarrollo de la metagenómica	<b>31-36</b>
1.2. Estudio del viroma respiratorio	<b>36-42</b>
1.2.1. Estudio del viroma pediátrico	<b>36-39</b>
1.2.2. El estudio del viroma como herramienta para el descubrimiento de nuevos virus	<b>39-42</b>
1.3. Aplicación de los métodos de secuenciación de nueva generación en la vigilancia de gripe	<b>43-48</b>
1.4. Secuenciación de adenovirus por métodos de secuenciación de nueva generación	<b>48-52</b>
<b>2. HIPOTESIS y OBJETIVOS</b>	<b>53-56</b>
<b>3. MATERIALES</b>	
3.1. Líneas celulares	<b>59</b>
3.2. Materiales de referencia utilizados como control virológico	<b>59-60</b>
3.3. Sujetos de estudio y muestras biológicas	<b>60-75</b>
3.3.1. Características de las muestras empleadas para el desarrollo de un protocolo para el estudio de viroma en las muestras clínicas respiratorias	<b>61-62</b>
3.3.2. Características de las muestras empleadas en el estudio de la aplicación de los métodos desarrollados en el estudio de la composición del viroma respiratorio y las relaciones ecológicas de sus integrantes.	<b>62-67</b>
3.3.3. Características de las muestras clínicas empleadas en el estudio de viroma como herramienta para descubrir posibles nuevos virus.	<b>67-71</b>

3.3.3.1.	Historial clínico “paciente 1”	<b>67-78</b>
3.3.3.2.	Historial clínico “paciente 2”	<b>68-69</b>
3.3.3.3.	Características de las muestras con infección respiratoria empleadas en un estudio poblacional sobre la presencia de un posible nuevo virus respiratorio en pacientes adultos y pediátricos	<b>69-71</b>
3.3.4.	Características de las muestras empleadas en el desarrollo y aplicación de técnicas NGS en el estudio de los virus gripales y su vigilancia epidemiológica	<b>71-73</b>
3.3.4.1.	Muestras empleadas en la estandarización de las técnicas NGS para el estudio de los virus gripales	<b>71-72</b>
3.3.4.2.	Historial clínico del “Paciente 1” y del “Paciente 2” en el estudio de pacientes pediátricos con gripe por WGS	<b>72-73</b>
3.3.5.	Muestras de campo empleadas en el estudio de adenovirus de diferentes especies de murciélagos de la Península Ibérica.	<b>74</b>
3.3.6.	Muestras de murciélagos empleadas en la puesta a punto de las técnicas de secuenciación de nueva generación para adenovirus	<b>74-75</b>
3.4.	Oligonucleótidos	<b>76-78</b>
3.4.1.	Oligonucleótidos publicados previamente a esta tesis	<b>76-77</b>
3.4.2.	Oligonucleótidos diseñados en esta tesis	<b>77-78</b>
3.4.3.	Sondas TaqMan empleadas en las diferentes técnicas de PCR tiempo real	<b>78</b>
3.5.	Reactivos	<b>79-82</b>
3.5.1.	Reactivos comerciales	<b>79-80</b>
3.5.2.	Kits comerciales	<b>80</b>
3.5.3.	Medios y reactivos de cultivo celular	<b>80-81</b>
3.5.4.	Equipo y aparatos	<b>81</b>
3.5.5.	Programas informáticos	<b>81-82</b>

3.5.6. Bases de datos	82
<b>4. MÉTODOS</b>	
4.1. Mantenimiento de líneas celulares	85
4.2. Aislamiento de virus en las líneas celulares VERO-E6, L132, Hep2 y MDCK-SIAT	85-86
4.3. Tratamiento de muestras clínicas	86
4.4. Extracción de ácidos nucleicos totales	87-88
4.4.1. Extracción manual por precipitación alcohólica de ácidos nucleicos totales	87
4.4.2. Extracción automática de ácidos nucleicos totales	87-88
4.5. Flujo de trabajo para la realización del estudio de viroma	88-95
4.5.1. Enriquecimiento viral por tres métodos	88-90
4.5.2. Amplificación independiente de secuencia con un único oligonucleótido cebador (SISPA)	90-92
4.5.2.1. SISPA para viroma ADN	90-91
4.5.2.2. SISPA para viroma ARN	91-92
4.6. Amplificación por círculo rodante (RCA)	93
4.7. Flujo de trabajo para la secuenciación de genoma completo del virus de la gripe	93-95
4.7.1. PCR multisegmento Gripe A	94
4.7.2. PCR multisegmento Gripe B	94
4.7.3. PCR multisegmento universal de gripe	94-95
4.8. Preparación de librerías <i>Nextera XT</i>	95-96
4.9. Método de enriquecimiento HyperCap Target Enrichment con preparación de librerías con el kit Kapa Hyper Plus	96-97
4.10. Análisis bioinformático de las secuencias obtenidas por NGS	98-100
4.10.1. Análisis bioinformático de las secuencias procedentes del estudio de viroma y de adenovirus de murciélagos	98-99
4.10.2. Análisis bioinformático de las muestras de gripe	99-100
4.11. Diseño de oligonucleótidos	100-101
4.12. PCRs tiempo real	101-103
4.12.1. PCR múltiples tiempo real BRQ y PCR múltiple tiempo	

real respiratorios 2	<b>101</b>
4.12.2. La PCR de identificación de Gripe tipos A, B y C	<b>102</b>
4.12.3. PCR MERS-CoV	<b>102-103</b>
4.13. PCR convencionales	<b>103-105</b>
4.13.1. PCR de estudio REP HRCiV	<b>103</b>
4.13.2. PCR de genoma completo de HRCiV	<b>103</b>
4.13.2.1. PCR de genoma completo 1	<b>103-104</b>
4.13.2.2. PCR de genoma completo 2	<b>104</b>
4.13.3. PCR de hexón para adenovirus	<b>104</b>
4.13.4. PCR ADN polimerasa para adenovirus	<b>104-105</b>
4.13.5. Visualización de productos de amplificación de las PCRs convencionales	<b>105</b>
4.14. Purificación del ADN y secuenciación	<b>105-108</b>
4.15. Análisis filogenéticos	<b>105-109</b>
4.15.1. Análisis filogenético de rinovirus detectados en pacientes pediátricos	<b>106</b>
4.15.2. Análisis filogenético de la familia <i>Redondoviridae</i>	<b>107</b>
4.15.3. Análisis filogenético de los virus de la gripe B linaje Yamagata de los pacientes pediátricos con cuadro gripal	<b>107</b>
4.15.4. Análisis filogenético del fragmento del hexón de adenovirus	<b>108</b>
4.15.5. Análisis filogenético del fragmento de ADN polimerasa de Adenovirus	<b>108-109</b>

## **5. RESULTADOS Y DISCUSION POR OBJETIVOS**

5.1. OBJETIVO I: Desarrollo del protocolo de trabajo para el estudio de viroma en muestras clínicas respiratorias	<b>113-116</b>
5.1.1. Estandarización de los métodos de enriquecimiento viral	<b>116-118</b>
5.1.2. Normalización del protocolo de amplificación independiente de secuencia con un cebador único para viroma ADN	<b>118-120</b>
5.1.3 Discusión	<b>121-124</b>

5.2. OBJETIVO II: Aplicación de los métodos desarrollados en el estudio de la composición del viroma respiratorio y las relaciones ecológicas de sus integrantes	<b>125-167</b>
5.2.1. Línea de trabajo 1: Estudio del viroma respiratorio de pacientes pediátricos. Presencia de RV y su posible relación con otros virus.	<b>125-126</b>
5.2.1.1. Virus identificados en pacientes con infección respiratoria empleando el flujo de trabajo para el estudio de viroma por NGS.	<b>127-131</b>
5.2.1.2. Caracterización de viroma del grupo control sanos	<b>131-133</b>
5.2.1.3. Caracterización de la presencia de RV en las muestras Pediátricas	<b>133-138</b>
5.2.1.4. Relación de RV con otros virus presentes en el tracto Respiratorio	<b>139-141</b>
5.2.1.5. Discusión	<b>141-149</b>
5.2.2. Línea de trabajo 2: El estudio de viroma como herramienta para descubrir nuevos virus	<b>149-151</b>
5.2.2.1. Caracterización virológica de pacientes con neumonía Grave	<b>151-153</b>
5.2.2.2. Secuenciación de virus circulares. RCA vs SISPA.	<b>153-154</b>
5.2.2.3. Caracterización del virus CRESS-DNA	<b>154-156</b>
5.2.2.4. Explorando la presencia del HRCiV en el tracto respiratorio	<b>156-158</b>
5.2.2.5. Papel de los virus HRCiV en la nueva familia <i>Redondoviridae</i>	<b>158-161</b>
5.2.2.6. Aislamiento viral del virus HRCiV en diferentes líneas celulares	<b>161-162</b>
5.2.2.7. Discusión	<b>162-167</b>
5.3. OBJETIVO III: desarrollo y aplicación de técnicas NGS en el estudio y referencia de los virus gripales y su vigilancia epidemiológica	<b>167-170</b>
5.3.1. Desarrollo de protocolos para el uso de WGS en el laboratorio	

de Referencia de Gripe	<b>170-172</b>
5.3.2. Amplificación de los ácidos nucleicos: selección de PCR multisegmento	<b>172</b>
5.3.3. Selección de la plataforma de secuenciación: <i>Illumina MiSeq</i> vs <i>MinION</i>	<b>172-176</b>
5.3.4. Análisis bioinformático de la WGS del virus de la gripe	<b>176-180</b>
5.3.5. Análisis del caso clínico de paciente pediátrica con gripe	<b>181-188</b>
5.3.6. Discusión	<b>188-192</b>
 5.4. OBJETIVO IV: Estudio de virus respiratorios ADN de gran tamaño y complejidad detectados en hospedadores no humanos: nuevos adenovirus en murciélagos ibéricos	 <b>193-194</b>
5.4.1. Estudio de la presencia de AdV en las muestras de murciélagos ibéricos	<b>194-196</b>
5.4.2. Análisis de las secuencias de hexón de AdV de murciélagos	<b>196-199</b>
5.4.3. Análisis de las secuencias de ADN polimerasa de AdV de Murciélagos	<b>199-203</b>
5.4.4. Estudio de los métodos de enriquecimiento para WGS de adenovirus	<b>203-205</b>
5.4.5. Infección de adenovirus de murciélago en células VERO E-6	<b>205-206</b>
5.4.6. Discusión	<b>206-212</b>
 6. <u>Epílogo</u>	 <b>211-218</b>
7. <u>Conclusiones</u>	<b>219-222</b>
8. <u>Bibliografía</u>	<b>223-239</b>
9. <u>Pósters y artículos</u>	<b>241-279</b>

## Abreviaturas y unidades de medida.

%: tanto por ciento

μL: microlitro

μM: micromolar

AB: aspirado bronquial

ABA: aspirado broncoalveolar

ABV: abreviatura

ADN: ácido desoxirribonucleico

AdV: adenovirus

ANF: aspirado nasofaríngeo

ARN: ácido ribonucleico

AT: aspirado traqueal

ATB: aspirado traqueobronquial

ATCC: Colección Americana de Cultivos Tipo "*American Type Culture Collection*"

BOV: bocavirus humano

cADN: ácido desoxirribonucleico complementario

CAP: proteína de la cápside

cel/μL: células por microlitro

cm<sup>2</sup>: centímetro cuadrado

COAST: *Childhood Origin of Asthma*

Contig: fragmento de secuencia sin ensamblar procedente de un análisis NGS

CoV: coronavirus



CRESS-DNA: virus ADN circulares que codifican para la proteína replicasa (*Circular rep-encoded single stranded DNA viruses*)

Ct: siglas en inglés para umbral de ciclo (*cycle threshold*)

ddNTPs: di-desoxinucleótidos trifosfato

dNTPs: desoxinucleótidos trifosfato

EAM: evento aparentemente mortal

EC: exudado conjuntival

ECACC: Colección Europea de Cultivos Celulares Controlados "*European Collection of Authenticated Cell Cultures*"

ECDC: Centro Europeo para el Control y Prevención de Enfermedades "*European Centre for Disease Prevention and Control*"

EF: exudado faríngeo

ENF: exudado nasofaríngeo

EOR: exudado orofaríngeo

EV: enterovirus

g/dL: gramos por decilitro

GB: Base de datos GenBank

GISAID: Base de datos "*Global Initiative on Sharing All Influenza*"

H: heces

HA: hemaglutinina

HRAPLV: virus respiratorio asociado a PSCV-5 (*Human respiratory-associated PSCV-5-like*)

HRCiV: virus respiratorio circular humano (*Human respiratory circular virus*)

HUH: His-aminoácidos hidrófobos-His

ICTV: Comité Internacional de Taxonomía de Virus. (*International Committee on Taxonomy of Viruses*)

IRA: infección respiratoria aguda

ITRS: infección del tracto respiratorio superior

Kb: kilobases

L/min: litro por minuto

LBA: lavado brocoalveolar.

LCR: líquido cefalorraquídeo

LN: lavado nasal

LNF: lavado nasofaríngeo

LQ AMN: líquido amniótico

LQ PER: líquido pericárdico

LQ pleural: líquido pleural

LQ SIN: líquido sinovial

M: matriz del virus de la gripe

M1: proteína de la matriz 1 del virus de la gripe

M2: proteína de superficie del virus de la gripe

MERS: síndrome respiratorio agudo del medio este

mg/dL: miligramos por decilitro

MNV: metapneumovirus

MO: médula ósea

N/A: no aplica

NA: neuraminidasa

NCBI: Centro Nacional de Información Biotecnológica (*National Center of Biotechnological Information*)

NGS: siglas en inglés de secuenciación de nueva generación (*Next Generation Sequencing*)

NP: nucleoproteína del virus de la gripe

NS: proteína no estructural del virus de la gripe

NS1: proteína no estructural 1 del virus de la gripe

NS2: proteína nuclear de exportación del virus de la gripe

nt: nucleótido

°C: grados centígrados

OMS: Organización Mundial de la Salud

PA: polimerasa ácida del virus de la gripe

pb: pares de bases

PB1: polimerasa básica 1 del virus de la gripe

PB2: polimerasa básica 2 del virus de la gripe

PCAD: enfermedad porcina asociada a circovirus (*Porcine Circovirus Associated Disease*)

PCR: reacción en cadena de la polimerasa

Per. Hosp: período de hospitalización

PM: marcador de peso molecular

PoSCV5: Virus circular porcino asociado a heces 5 (*Porcine stool-associated circular virus 5*)

RCA: siglas en inglés de amplificación por círculo rodante (*Rolling Circle Amplification*)

Ref: referencia bibliográfica

REP: proteína replicasa

RFU: siglas en inglés para unidad de fluorescencia relativa (*relative fluorescence units*)

RT: retrotranscripción o transcripción inversa

RuV: Rubulavirus humano

RV: rinovirus

SARS: síndrome respiratorio agudo severo (*Severe Acute Respiratory Syndrome*)

SECEMU: Sociedad Española para la Conservación y el Estudio de los Murciélagos

SF3: helicasa de la superfamilia 3

SISPA: siglas en inglés de amplificación independiente de secuencia con un cebador único (*Single Independent Sequence Primer Amplification*)

SP: esputo

TTV: virus Torque Teno

UCI: unidad de cuidados intensivos

VPI: virus parainfluenza

VRS: virus respiratorio sincitial

WGS: secuenciación de genoma completo (*Whole genome sequencing*)

## Apéndice de tablas y figuras

### TABLAS

Tabla 1: Relación de virus control empleados en las diferentes técnicas utilizadas

Tabla 2: Tipo de muestras recibidas en el Laboratorio de Virus Respiratorios y el código de identificación de dichas muestras

Tabla 3: Identificación y caracterización virológica de virus respiratorios obtenidos por métodos moleculares en las muestras empleadas en el objetivo I

Tabla 4: Presentación clínica de los 153 pacientes pediátricos empleados en el estudio del viroma respiratorio. La edad está expresada en meses y el período de hospitalización está expresado en días

Tabla 5: Relación de muestras tomadas y día de toma de la muestra de la paciente 1

Tabla 6: Identificación y tipo de las muestras del paciente 2

Tabla 7: Aislados de gripe, tipo y subtipo de los virus y tipo de muestra de la que proceden los virus empleados en las diferentes estandarizaciones de los protocolos de WGS de gripe

Tabla 8: Identificación y procedencia de las muestras de la Paciente 1 para el análisis WGS del virus de la gripe

Tabla 9: Identificación y tipo de muestra del “Paciente 2” para el análisis WGS del virus de la gripe

Tabla 10: Familia, especie, tipo de muestra y año de captura de los murciélagos empleados en el estudio de AdVs de murciélagos de la Península Ibérica

Tabla 11: Identificación, tipo de muestra y especie de los murciélagos empleados en el estudio para WGS de adenovirus

Tabla 12: Secuencia, nombre y PCR en la que se emplean estos oligonucleótidos cebadores procedentes de trabajos previamente publicados

Tabla 13: Nombre, secuencia del oligonucleótido cebador empleado y tamaño del producto de amplificación de las PCRs diseñadas en esta tesis

Tabla 14: Sondas utilizadas en las técnicas de PCR múltiple tiempo real empleadas en esta tesis

Tabla 15: Identificación de la muestra analizada y tamaño de secuencia de los virus obtenidos en muestra respiratoria enriquecida y no enriquecida

Tabla 16: Identificación de las muestras analizadas y tamaño de secuencia de los virus obtenidos en muestra enriquecida por 3 métodos y enriquecida mediante captura por sondas

Tabla 17: Resultados de la prueba de la comparación entre los virus detectados en viroma ADN y viroma ARN y su mezcla equimolar

Tabla 18: Patógenos respiratorios virales detectados tras el análisis de las muestras respiratorias con IRA

Tabla 19: Número total de fragmentos y lecturas de los virus detectados en las muestras de pacientes con IRA pero sin identificación de patógeno viral

Tabla 20: Número total de fragmentos y lecturas de los virus detectados en las muestras de pacientes control

Tabla 21: Línea temporal de la toma de muestras positivas para RV y los tipos de RV detectados en ellas

Tabla 22: Porcentaje de homología entre los diferentes rinovirus secuencias por NGS de los pacientes con muestras de seguimiento

Tabla 23: Método de detección, tipo de RV detectado y diagnóstico de los pacientes a los que le fue detectado TTV

Tabla 24: Número de lecturas de virales de los extractos ADN y ARN tras el análisis por secuenciación masiva y los resultados obtenidos en la PCR de estudio de las muestras clínicas de la paciente 1

Tabla 25: Número de lecturas coincidentes con el virus HRPLAV obtenidas por secuenciación masiva y el resultado de la PCR de estudio realizada a las muestras respiratorias del paciente 2

Tabla 26: Identificación, detección y tamaño de secuencia obtenidos empleando los métodos SISPA ADN y RCA

Tabla 27: Identificación, diagnóstico y virus respiratorios encontrados en los pacientes positivos a HRCiV del grupo 1. Los pacientes que solo presentaron infección con HRCiV están destacados en negrita

Tabla 28: Diagnóstico y presencia de otros virus en los pacientes positivos del grupo 2

Tabla 29: Fecha de toma de muestra, edad y resultados de la PCR REP HRCiV de los hermanos pediátricos de estudio

Tabla 30: Porcentaje de distancia en pares entre los virus encontrados en el estudio de caracterización y la paciente 1

Tabla 31: Umbrales de ciclo de los virus de la gripe analizados obtenidos tras la PCR tiempo real Gripe A, B y C en muestras enriquecidas viralmente y sin enriquecimiento viral

Tabla 32: Resultados obtenidos del tamaño y la homología de las secuencias consenso obtenidas

Tabla 33: Segmentos del genoma obtenido de los diferentes virus analizados con las secuencias de referencia para el ensamblaje empleadas en *INSAFLU* empleando los secuenciadores *MinION* *Illumina MiSeq*

Tabla 34: Secuencias de virus obtenidas por WGS y número de identificación de los virus depositados en la base de datos *GISAID*

Tabla 35: Resultados del análisis de las 12 muestras procedentes de la paciente pediátrica tras realizarles las PCRs tiempo real Gripe A, B y C, múltiple BRQ, múltiple Respiratorios 2 y las PCRs convencionales multisegmento A y B

Tabla 36: Resultados del análisis de las muestras procedentes del paciente 2 tras realizarles las PCRs tiempo real Gripe A, B y C, múltiple BRQ, múltiple Respiratorios 2 y las PCRs convencionales multisegmento A y B

Tabla 37: Especies de murciélago estudiadas, resultados del estudio de AdV en EOR y H, año de captura y número de identificación de *GenBank*

Tabla 38: Clasificación de los mastadenovirus de murciélago tras el análisis de distancia entre pares de la secuencia de aminoácidos obtenida de las secuencias del fragmento de ADN polimerasa

Tabla 39: Tamaño de secuencia obtenido tras el análisis bioinformático de las secuencias obtenidas tras la aplicación de los dos métodos de enriquecimiento de secuencias comparados

## FIGURAS

Figura 1: Retos en la utilización de la metagenómica clínica

Figura 2: Esquema de los distintos tipos muestra respiratoria empleada procedentes de los pacientes pediátricos para el estudio de viroma.

Figura 3: Esquema de los grupos de pacientes empleados en el estudio de presencia del nuevo virus en el tracto respiratorio y del origen de las muestras tomadas en cada grupo de análisis

Figura 4: Abundancia del tipo de muestra respiratoria empleada en el estudio de la presencia del nuevo virus respiratorio en adultos

Figura 5: Distribución de la presencia de virus respiratorios en las muestras respiratorias pediátricas empleadas en el grupo 2 del estudio de presencia del nuevo virus en el tracto respiratorio

Figura 6: Mapa de la distribución geográfica de los lugares de captura de los murciélagos analizados

Figura 7: Flujo de trabajo para el estudio de viroma

Figura 8: Imagen del degradado característico de SISPA ADN

Figura 9: Flujo de trabajo del protocolo *Nextera XT* para la preparación de librerías

Figura 10: Flujo de trabajo del protocolo de preparación de librerías *Kapa Hyper Plus* y de enriquecimiento con el panel de captura por sondas *HyperCap Target Enrichment*

Figura 11: Flujo de trabajo sobre las muestras respiratorias pediátricas para comparar el número de lecturas virales entre las muestras enriquecidas y no enriquecidas

Figura 12: Flujo de trabajo sobre las muestras pediátricas para comparar el número de lecturas virales entre las muestras enriquecidas por 3 métodos y las enriquecidas por captura por sondas

Figura 13: Flujo de trabajo empleado en el estudio de la aplicación de la enzima *Klenow exopolimerasa* en uno o dos pasos



Figura 14: Representación gráfica del número de lecturas virales obtenidas tras el análisis de viroma ADN y ARN de las muestras enriquecidas y no enriquecidas

Figura 15: Número de lecturas virales obtenidas con la 2<sup>a</sup> incubación de la enzima *Klenow exopolimerasa* y sin la 2<sup>a</sup> incubación

Figura 16: Método de enriquecimiento, de preparación de librerías y muestras empleadas en caracterizar el viroma de pacientes sanos (40+21 ANF), pacientes con IRA (57 ANF) y en el estudio de la presencia de RV en muestras de seguimiento (35 ANF)

Figura 17: Abundancia relativa de los fragmentos obtenidos tras el análisis de las muestras respiratorias de los pacientes pediátricos con IRA

Figura 18: Identificación taxonómica de los fragmentos obtenidos tras el análisis de las muestras respiratorias de los pacientes pediátricos con IRA

Figura 19: Abundancia relativa de los fragmentos obtenidos tras el análisis de las muestras respiratorias de los pacientes pediátricos control

Figura 20: Identificación taxonómica de los fragmentos obtenidos tras el análisis de las muestras respiratorias de los pacientes pediátricos control sanos

Figura 21: Identificación taxonómica de los virus detectados tras el análisis de las muestras respiratorias de seguimiento de los pacientes pediátricos con RV

Figura 22: Distribución del tipo de RV detectados en las muestras de seguimiento de pacientes pediátricos

Figura 23: Análisis filogenético de los gen VP1 de los RV detectados en esta tesis y rinoviruse de referencia. El análisis fue realizado con el programa *MEGA 5.2* usando el método *Neighbor-Joining* con un modelo *Tamura-Nei*, con un respaldo estadístico de 5000 repeticiones

Figura 24: Identificación taxonómica de los virus detectados tras el análisis de las muestras respiratorias de seguimiento de los pacientes control sanos

Figura 25: Representación gráfica del genoma circular del virus HRCiV

Figura 26: Motivos conservados RCR y Walker de la proteína REP de los virus HRCiV 15040 y HRCiV 15278

Figura 27: Análisis filogenético de los virus HRCiV detectados en esta tesis y los virus de la familia *Redondoviridae* publicados hasta el momento. El análisis fue realizado con el programa MEGA 5.2 usando el método *Maximum Likelihood* con un modelo *Hasegawa-Kishino-Yano*, con un respaldo estadístico de 5000 repeticiones

Figura 28: Flujo de trabajo para WGS del virus de gripe empleado en las muestras de la paciente 1

Figura 29: Resultados comparativos entre muestras enriquecidas y no enriquecidas. Gel de agarosa al 1% comparando los virus de Gripe A (H3N2) “A/Navarra/3251/2016” y “A/CastillaLaMancha/3399/2016” tras realizarse el protocolo de enriquecimiento con 3 métodos y sin la realización de éste

Figura 30: Gel de agarosa al 1% que muestra los resultados obtenidos en la amplificación por PCR multisegmento A (llave azul), B (llave naranja) y la PCR multisegmento universal (llaves amarillas) usando la dilución -2 y -3 de los virus A (H1N1 pdm) y B (linaje Yamagata). CAL: “A/California/07/2009”. PHU: B/Phuket/3073/2013

Figura 31: Fragmento del alineamiento de las secuencias de aminoácidos de las secuencias consenso de la HA de gripe obtenidas en la secuenciación con *MiSeq*, *MinION* y el método Sanger. El alineamiento fue realizado con el programa *Clustal X2* y la traducción a aminoácidos del alineamiento fue realizado con el software *MEGA 5.2*

Figura 32: Gráfico generado por *INSaFLU* que muestra la cobertura y secuenciación de los ocho segmentos de gripe de los 3 virus de gripe A analizados. 172444: “A/Galicia/2444/2017”. 172467: “A/Galicia/2467/2017”. 172479: “A/Baleares/2479/2017”

Figura 33: Gráfico generado por *INSaFLU* que muestra la cobertura y secuenciación de los ocho segmentos de gripe de los 3 virus de gripe B analizados empleando como secuencia de referencia “B/Norway/2409/2017”. 172407: “B/Galicia/2407/2017”. 172406: “B/Galicia/2406/2017”. 172439: “B/CastillaLaMancha/2439/2017”

Figura 34: Gráfico generado por *INSaFLU* que muestra la cobertura y secuenciación de los ocho segmentos de gripe de los 3 virus de gripe B analizados empleando como secuencia de referencia “B/Brisbane/60/2008”. 172407: “B/Galicia/2407/2017”. 172406: “B/Galicia/2406/2017”. 172439: “B/CastillaLaMancha/2439/2017”

Figura 35: Gráfico generado por *IRMA* de la distribución de lecturas de los 8 segmentos de los virus de la gripe A “A/Galicia/2444/2017”, “A/Galicia/2467/2017”, “A/Baleares/2479/2017”

Figura 36: Gráfico generado por *IRMA* de la distribución de lecturas de los 8 segmentos de los virus de la gripe B, “B/Galicia/2406/2017”, “B/Galicia/2407/201” y “B/CastillaLaMancha/2439/2017”

Figura 37: Diagrama de cambio de nucleótidos y las frecuencias en la que se producen en el segmento M del virus “B/Palencia/1434/2018”. Gráfico generado en la plataforma *IRMA*

Figura 38: Alineamiento realizado con el programa *MACAW* entre los fragmentos de M gripe A detectados en la muestras 181434 y la gripe A H3N2 “A/Singapore/INFIMH-16-0019/2016”

Figura 39: Análisis filogenético realizado por el modelo *Hasegawa-Kishino-Yano* del método *Maximum Likelihood* con un respaldo estadístico de 5000 repeticiones. En el que se muestra la agrupación del fragmento obtenido con el grupo de la gripe circulante “B/Norway/2409/2017” del linaje Victoria en vez de con el grupo de “B/Phuket/3073/2013” del linaje Yamagata con el que se agrupaba en el resto de segmentos

Figura 40: Alineamiento realizado con el programa *MACAW* entre los fragmentos de M gripe A detectados en la muestra 181434 tras los dos métodos WGS realizados y la gripe A H3N2 “A/Singapore/INFIMH-16-0019/2016”

Figura 41: Análisis filogenético realizado por el modelo *Hasegawa-Kishino-Yano* del método *Maximum Likelihood* con un respaldo estadístico de 5000 repeticiones. En el que se muestra el análisis filogenético de los segmentos PB2, PB1, PA, HA, M y NA de los virus “B/Palencia/1434/2018” y “B/Palencia/1434B/2018” junto con otros virus de la gripe B

Figura 42: Árbol filogenético basado en el análisis del fragmento del gen de hexón. El árbol fue estimado con el programa *MEGA 5.2* usando el método *Neighbor-Joining* con el modelo *Tamura 3 parámetros*. El soporte estadístico fue de 5000 repeticiones

Figura 43: Árbol filogenético basado en el análisis del fragmento del gen de la ADN polimerasa. El árbol fue estimado con el programa *MEGA 5.2* usando el método *Neighbor-Joining* con el modelo *Kimura 2 parámetros*. El soporte estadístico fue de 5000 repeticiones

Figura 44: Distribución del número de lecturas obtenidas en los 7 EOR empleados tras el enriquecimiento por 3 métodos y la preparación de librerías con el kit *Nextera XT*.

Figura 45: Distribución del número de lecturas obtenidas en los 7 EOR empleados tras el enriquecimiento mediante captura por sondas y la preparación de librerías con el kit *HyperCap Target Enrichment*

Figura 46: Alineamiento realizado con el programa *MACAW* entre las secuencias del mastadenovirus de murciélago de referencia BtAdV-2 y los AdV “Pku15 150518” y “Ppy15 181106” obtenidas tras el enriquecimiento mediante captura por sondas

Figura 47: Relación entre las tecnologías estandarizadas para cada estudio definido en los objetivos de trabajo de esta tesis y su influencia. Las líneas rojas representan la influencia que ha tenido el desarrollo del método de amplificación independiente de secuencia en la secuenciación de gripe y de adenovirus. Las líneas verdes representan la influencia que la comparación entre secuenciadores para muestras de gripe ha tenido en la secuenciación de las muestras del estudio de viroma y de adenovirus. Las líneas moradas representan la influencia de los diferentes métodos de enriquecimiento en muestras de murciélagos y muestras respiratorias humanas empleadas en el estudio del viroma y de secuenciación de gripe



# Introducción



## INTRODUCCIÓN

### 1.1 Desarrollo de la metagenómica clínica

La virología clínica es un amplio campo que comprende desde la identificación de patógenos para el diseño de estrategias de tratamiento de la enfermedad hasta la vigilancia de patógenos importantes en salud pública y se basa en el diagnóstico y la investigación virológica.

El diagnóstico y la investigación tradicional de los virus han consistido, fundamentalmente, en una serie de pruebas que generalmente son específicas para cada virus y están basadas en el crecimiento de los virus en cultivos celulares (aislamiento), la detección de los anticuerpos específicos frente a ellos (serología) o la identificación molecular por la reacción en cadena de la polimerasa (PCR) [1].

En la actualidad, los resultados obtenidos por PCR y su posterior secuenciación usando el método Sanger son los métodos más comunes en los laboratorios clínicos y de salud pública. Estos métodos son determinantes a la hora de establecer un diagnóstico y conocer las características de un virus detectado. Es por ello que la mejora del rendimiento de cualquiera de estos abordajes conlleva una mejora de los resultados obtenidos.

La secuenciación Sanger [2] forma parte de la primera generación de tecnologías para la secuenciación del ácido desoxirribonucleico (ADN). El método consiste en la síntesis de una cadena simple utilizando para la reacción de secuenciación los cuatro desoxinucleótidos trifosfato (dNTP) que conforman el ADN, una polimerasa ADN y cuatro di-desoxinucleótidos trifosfato (ddNTPs). Estos ddNTPs marcados con fluoróforos interrumpen la adición del nucleótido consecutivo generando diferentes tamaños de fragmentos que serán analizados por electroforesis capilar en un gel de poliacrilamida. Este método usado durante más de tres décadas sigue vigente actualmente y a pesar de su utilidad presenta limitaciones en el número de secuencias que se generan y en la capacidad de reflejar la diversidad viral en la muestra.

Estas limitaciones pueden ser mejoradas gracias a los métodos de secuenciación de nueva generación (NGS). La aplicación clínica de la metagenómica tiene sus raíces en el comienzo del nuevo siglo con el desarrollo de los *microarrays* [3], que entre sus éxitos cuenta con el descubrimiento del nuevo coronavirus asociado al SARS [4]. Sin embargo, el gran salto cualitativo y cuantitativo de la metagenómica clínica viene de la mano de la tecnología NGS y su capacidad de generar en una sola carrera millones de secuencias permitiendo el análisis de genomas



completos y detectando la presencia de diferentes virus, bacterias y hongos en una misma muestra.

Hasta el momento varias plataformas NGS han sido desarrolladas y son comúnmente usadas como *Roche 454 GS FLX*, *Illumina (MiSeq y HiSeq)*, *Ion Torrent/IonProton*, sistema *SOLiD 5500* y *Oxford Nanopore* [5].

El uso del secuenciador *Illumina MiSeq* y los kits de preparación de librerías de la misma casa comercial (*Illumina*) están siendo mayoritariamente empleados debido a su precio competitivo y su capacidad de generar fragmentos largos fiables [6]. Este método está basado en terminadores reversibles que detectan bases únicas a medida que se incorporan masivamente a cadenas de ADN paralelas. Los colorantes fluorescentes de los terminadores se fotografían con la adición de cada dNTP y a continuación se eliminan para permitir la incorporación de la siguiente base. Las llamadas de bases se realizan directamente a partir de mediciones de intensidad de señal durante cada ciclo [5].

El desarrollo de estas tecnologías no se ha estancado a pesar de contar con equipos fiables y cada vez más asequibles. Una de las nuevas alternativas para la secuenciación masiva es el dispositivo de bolsillo *MinION* desarrollado por la casa comercial *Oxford Nanopore Technologies*. Este dispositivo permite secuenciar y realizar análisis bioinformáticos como son el subtipado de los virus de gripe o el ensamblado de secuencias para obtener virus completos [7]. Su tamaño reducido y su capacidad de análisis le hacen ser una buena alternativa en caso de una posible pandemia gripal o para la vigilancia de gripe en lugares remotos. Su método de secuenciación se basa en la identificación de bases midiendo cambios en la conductividad eléctrica de un poro biológico al paso del ADN [8].

A pesar de la ventana de conocimiento que se abre con estas nuevas tecnologías las dificultades son numerosas y se presentan en cada paso del proceso, como se ilustra en la Figura 1. Los obstáculos más importantes son: 1. la necesidad de implementar un tratamiento y una extracción de ácidos nucleicos de la muestra a estudiar que sean los más adecuados para aumentar la proporción de virus; 2. el diseño de nuevas técnicas para la amplificación del genoma viral completo; 3. la elección del método más preciso para la preparación de librerías y 4. el manejo de la plataforma para la realización de los análisis bioinformáticos para el estudio de las secuencias obtenidas [1].

Comparada con otros microorganismos, la abundancia relativa de virus en la muestra clínica a analizar es un punto crítico para la detección de genomas completos en metagenómica viral. Una carga viral alta aumenta la probabilidad de que toda la variedad viral sea representada en su totalidad de manera fiable y para ello se han desarrollado diferentes protocolos de enriquecimiento viral.

La aplicación de técnicas como la centrifugación, la filtración, el tratamiento enzimático o la combinación de las tres permite de manera simple y barata aumentar la proporción de virus en la muestra a secuenciar [9]. Otro de los sistemas disponibles de enriquecimiento es el diseño de sondas de captura basadas en secuencias representativas de virus que aumentan la cantidad de éstas frente a las secuencias de fondo procedentes del hospedador u otros microorganismos. Estas técnicas de enriquecimiento no funcionan de igual manera para todos los tipos de virus [9], por ello un estudio previo sobre la preparación y tipo de extracción que se va a realizar sobre un tipo de muestra biológica concreta [10] es de vital importancia para determinar la mejor estrategia y obtener los mejores resultados posibles usando NGS.

Además del enriquecimiento, otros métodos pueden ser implementados para aumentar la proporción de secuencias de nuestro interés en la muestra clínica. En este trabajo hemos desarrollado tres métodos diferentes aplicados a los distintos estudios realizados: 1. el método de amplificación independiente de secuencia con un cebador único (SISPA) para el estudio del viroma; 2. la amplificación por círculo rodante (RCA) para la caracterización de genomas circulares, y 3. el uso de PCRs multisegmento para la amplificación simultánea de los ocho segmentos del genoma del virus de la gripe.

El método SISPA [10] es una técnica que permite la amplificación de virus muy diversos y la posibilidad de amplificar varios genomas virales completos presentes en la muestra de manera simultánea. Este método inicialmente desarrollado por Reyes y Kim [11] ha sufrido diversas modificaciones [12,13] pero fundamentalmente implica la unión de un nucleótido cebador aleatorio con un hexámero de “N” a una molécula de ADN ó ADN complementario (ADNc) a la que se le realizarán varios ciclos de amplificación para aumentar la cantidad de ADN ó ADNc disponible en la muestra.

En el caso de los virus circulares uno de los procesos más simples y eficientes es la amplificación en círculo rodante ó RCA según sus siglas en inglés [14]. Este protocolo es un proceso enzimático

isotérmico [15] donde se usa la ADN polimerasa  $\phi$ 29 para producir una hebra de ADN de cadena simple complementaria a la hebra molde. Este protocolo necesita únicamente 4 elementos: la ADN polimerasa con su tampón homólogo, unos nucleótidos cebadores aleatorios, una hebra molde circular y dNTPS.

Finalmente, las amplificaciones previas a la secuenciación pueden ser también PCRs dirigidas y específicas que aumentarán el número de secuencias de ADN ó ADNc disponibles del virus de nuestro interés. Un ejemplo de enorme interés es el caso de los virus de la gripe para los que han sido desarrollados varios conjuntos de oligonucleótidos cebadores diseñados para amplificar cada uno de los segmentos genómicos de los virus completos de gripe A [16], de gripe B [17] o ambos tipos con un solo set de oligonucleótidos cebadores [18].

A pesar de todos estos avances en amplificación y secuenciación el reto más importante para los laboratorios de salud pública sigue siendo el análisis de las secuencias procedentes de NGS. La obtención de millones de secuencias requieren análisis realizados por personal con conocimientos en bioinformática y una capacidad computacional a gran escala [19]. Actualmente las herramientas de uso podrían distinguirse entre plataformas de análisis basadas en línea de comandos que requieren personal especializado tanto en su diseño como en su ejecución, plataformas comerciales cerradas (*CLC Genomics Workbench de Qiagen, Geneious, Bionumerics de Applied Maths o Ridom SeqSpher+ de Ridom Bioinformatics*) [20–23] o plataformas web abiertas y gratuitas que incluyen la extracción de las secuencias de bacterias y del genoma del hospedador y además, realizan el ensamblaje *de novo* de las secuencias virales como la plataforma *Virus TAP* ó *INSaFLU* [24,25].

La implementación de esta tecnología es un proceso en constante cambio para estandarizar cada protocolo al objetivo concreto deseado, por ello cada procedimiento debe ser específico y susceptible de mejora constante debido a la rápida evolución de los elementos básicos de esta tecnología [26]. Sin embargo cada mejora específica de un protocolo puede aportar información para mejorar otros, creando un flujo de trabajo que, a pesar de estar dirigido a la consecución de un objetivo concreto, alimenta y se alimenta de otras estrategias. Esta es la razón por la que es un método de trabajo muy provechoso para la investigación y el diagnóstico de los virus respiratorios debido a que puede ser empleado pese a la variedad taxonómica y filogenética viral, la diversidad de los pacientes que infectan y las etiologías específicas.



**Figura 1:** Retos en la utilización de la metagenómica clínica. Durante cada paso del proceso varios puntos son críticos para la implementación de un protocolo de trabajo en metagenómica. Imagen tomada de (*Chiu y Miller, 2019*).

En este trabajo se ha desarrollado el estudio sobre viroma respiratorio en el que se ha intentado dar respuesta a cada reto (Figura 1) desarrollando protocolos ajustados a las muestras respiratorias y a los virus que buscamos. En esta tesis se ha aprovechado la oportunidad que ofrecen estas nuevas tecnologías de secuenciación para adaptarlas a la investigación, humana y animal, y a la vigilancia de un virus tan significativo para la salud pública como es la gripe.

## 1.2 Estudio del viroma respiratorio

Se define como viroma humano al conjunto de virus patógenos, virus comensales con capacidad de replicar en células humanas de pacientes sanos, bacteriófagos y elementos retrovirales integrados en el cromosoma humano [27,28]

Tradicionalmente el estudio virológico se ha centrado en los patógenos debido a su asociación con enfermedades y a su impacto en salud pública. Como hemos mencionado previamente, los métodos utilizados para la caracterización viral han sido en los últimos tiempos métodos moleculares muy sensibles y específicos por lo que hasta el desarrollo de las nuevas técnicas NGS el estudio del viroma estaba muy limitado.

En el caso del viroma respiratorio estos nuevos métodos han supuesto un nuevo impulso ya que han permitido la posibilidad de estudiar la comunidad viral y la identificación de nuevos patógenos de una muestra de forma directa sin necesidad de cultivos, clonaje o cualquier conocimiento previo de la muestra como la caracterización de los virus presentes en la muestra.

Describir el viroma respiratorio y conocer la relevancia de los virus detectados es una tarea importante para conocer en profundidad la infección respiratoria pudiendo cambiar el paradigma de lo que consideramos salud o enfermedad en el tracto respiratorio.

### **1.2.1 Estudio del viroma pediátrico**

La infección respiratoria aguda (IRA) es responsable de cerca de 3 millones de muertes anuales en el mundo, fundamentalmente en niños y ancianos, sin una reducción sustancial en la mortalidad estimada durante las dos pasadas décadas [29]. Esta patología es responsable de una mortalidad del 1-3% en niños menores de 5 años en países industrializados y del 10-25% en países en vías de desarrollo [30]. Muchas de estas IRAs son de origen vírico siendo una de las principales causas de hospitalización [31].

Los agentes etiológicos relacionados con IRAs virales son los virus de la gripe, los virus parainfluenza (VPI), el virus respiratorio sincitial (VRS), enterovirus (EV), coronavirus (CoV), el metapneumovirus humano (MNV), bocavirus (BOV), adenovirus (AdV) y los rinovirus (RV).

Las manifestaciones de las infecciones víricas son muy variables, con un espectro clínico que incluye desde infecciones leves, que pueden ser atendidas de forma ambulatoria, a formas más graves que precisan hospitalización de duración variable. Además, un único agente puede dar lugar a cuadros clínicos muy distintos, mientras que varios agentes infecciosos pueden dar lugar a varios síndromes semejantes, no diferenciables clínicamente [32].

Las infecciones virales respiratorias afectan a niños de todas las edades, si bien son más frecuentes en niños menores de 5 años, y especialmente prevalentes en menores de 2 años. La

etiología de las infecciones respiratorias ha sido bien estudiada en niños hospitalizados, conociéndose en la actualidad los agentes más frecuentemente implicados en las infecciones moderadas o graves que precisan ingreso. Así, en los niños por debajo de 2 años los virus más frecuentes son, en primer lugar el VRS, alcanzando cerca del 50% de las infecciones que requieren de hospitalización, y en segundo lugar los RV, en torno al 30%. Los cuadros clínicos a los que se asocian estos virus son bronquiolitis y episodios de sibilancias recurrentes como causas más frecuentes de hospitalización, seguidos de neumonías, laringitis, síndromes febriles o gripales y cuadros catarrales en los lactantes más pequeños [33–36].

En los niños mayores de 2 años, los virus más frecuentemente implicados en la hospitalización son el grupo de los RV, seguidos por diferentes tipos de AdV, el BOV y en menor proporción otros virus como los VPI, virus de la gripe y VRS. Los cuadros clínicos asociados son fundamentalmente crisis asmáticas y neumonías [36].

En los niños con patología respiratoria ambulatoria los estudios son menos frecuentes, aunque también existen datos. Los virus más frecuentes en este grupo de pacientes son los RV, seguidos del VRS, y en menor porcentaje los CoV, VPI, MNV y BOV. Los cuadros clínicos en pacientes extrahospitalarios suelen ser cuadros catarrales de vías altas, bronquiolitis, sibilancias recurrentes y laringitis [37–40].

Los RV son virus pertenecientes al género *Enterovirus* de la familia *Picornaviridae*. Presentan un ARN de cadena simple con polaridad positiva y genomas con un tamaño comprendido entre 7.2 y 8.5 kilobases (kb). El primer RV fue descrito en 1953, y en 1987 ya se habían descrito 100 serotipos distintos del inicial, mostrando la enorme variabilidad de estos virus [41,42]. La secuenciación completa y parcial (fundamentalmente las regiones codificantes VP4/VP2 y VP1) de cepas y virus presentes en muestras clínicas reveló la existencia de dos grandes genotipos o clados de RV, que correlacionaba principalmente con diferente sensibilidad a antivirales, dando lugar a la división en dos tipos de rinovirus: RVA y RVB [43,44]. El desarrollo de técnicas moleculares, más sensibles para la detección de RV en muestras clínicas, llevó al descubrimiento de una nueva especie de rinovirus, el denominado rinovirus C [45]. El genoma de estos nuevos virus presentaba menos del 70% de identidad en aminoácidos en las proteínas P1, 2C y 3CD, como indica el criterio de demarcación de especie del género *Enterovirus*, y se caracterizaron por no poder ser cultivados en las líneas celulares más comúnmente empleadas para el crecimiento de los virus respiratorios [42]. Los RVC han sido detectados en las vías respiratorias superiores e

inferiores y en algunos estudios se han identificado como el tipo de virus dominante en más de la mitad de las infecciones respiratorias pediátricas diagnosticadas en todo el mundo [42].

Los RV han sido relacionados con sibilancias infantiles y asma [46], especialmente los RVA y RVC que son considerados 7 veces más propensos que RVB a causar enfermedad [47]. De acuerdo a su potencial relación con la aparición de sibilancias o asma, estudios de cohortes previos aislaron el virus en el 90% de los pacientes con brotes de asma [48] y probaron su conexión con la hospitalización por IRA [49]. Uno de los estudios más completos sobre asma e infección por RV fue el trabajo conocido como “*Childhood Origin of Asthma (COAST)*” que mostró que haber padecido una infección por RV era el factor de riesgo más significativo para presentar sibilancias permanentes en niños de 3 años [50]. Además, presentar sibilancias causadas por RV no solo es un factor de riesgo importante para tener asma en la infancia sino que también se relaciona con presencia de asma en adolescentes [51].

El virus Torque Teno (TTV) es uno de los principales representantes de los virus comensales. Este virus, que pertenece a la familia *Anelloviridae*, fue descubierto en 1997 [52]. Posee un genoma circular de ADN de cadena simple de un tamaño de 2.2 a 3.7 kb. Son virus caracterizados por su ubicuidad en todo tipo de tejidos [53], y por presentar infección persistente durante toda la vida, por lo que a pesar de no ser considerados patógenos esta presencia permanente puede tener influencia sobre el sistema inmune. Trabajos previos sobre la viremia del TTV del genogrupo 4 mostraron niveles más altos de viremia en pacientes pediátricos con bronconeumonía frente a pacientes con una infección aguda más leve [54]. Acerca del nexo del TTV con el sistema inmune se ha descrito una asociación positiva entre la cantidad de TTV presente en nariz y los niveles de proteína catiónica eosinófila, siendo esta proteína un marcador de inflamación bronquial especialmente elevado en niños asmáticos [55]. A pesar de estos estudios sobre la posible relación de los TTV con la infección respiratoria, y su relación con el sistema inmune, el rol concreto de este virus en el desarrollo de IRA o asma continúa por ser descrito. Dado que trabajos previos han detectado coinfecciones de rinovirus y Torque Teno virus en pacientes adultos con IRA [56] consideramos que el conocimiento profundo de la asociación entre estos dos virus puede ayudar a definir mejor las causas de la IRA y del asma y su potencial origen en las infecciones infantiles

### 1.2.2 El estudio del viroma como herramienta para el descubrimiento de nuevos virus

El desarrollo de la metagenómica viral ha demostrado ser una herramienta poderosa para conocer la diversidad viral de diferentes ecosistemas [57]. El uso de estos métodos ha potenciado el descubrimiento de numerosos virus en muestras ambientales y animales potencialmente relacionados con la salud pública [58–63].

En el caso del estudio de muestras de procedencia humana las técnicas empleadas para el estudio de viroma han descrito recientemente más virus nuevos que los caracterizados por técnicas clásicas, dibujando un nuevo escenario en el que conocemos nuevas especies virales pero no su impacto desde el punto de vista de la infección asociada.

Entre estos nuevos virus descritos gracias a la secuenciación de nueva generación se encuentran los virus ADN circulares de cadena simple, conocidos por su acrónimo en inglés como virus CRESS-DNA, que codifican para la proteína replicasa (REP). Este término fue acuñado en 2012 y describe al grupo de virus encapsulados más pequeños conocidos que infectan a eucariotas y que parecen tener un ancestro común. Dentro de este grupo en los extremos del espectro virológico, 11 de los 13 virus de cadena simple reconocidos por el comité internacional de taxonomía de virus (ICTV) son de genoma circular. Siete de estos virus infectan a organismos eucariotas y de éstos únicamente los miembros de la familia *Anelloviridae* no poseen una REP homóloga [64].

Estos virus han sido descritos en un amplísimo rango de distribución y hospedadores. La causa de su diversidad puede deberse a su potencial evolutivo. Los virus CRESS-DNA se caracterizan por su predisposición a la recombinación [65,66] y su alta tasa de sustitución (1,2 por  $10^{-3}$  sustituciones por sitio al año), similar al número de sustituciones encontradas en los virus ARN, mostrando una rápida evolución en comparación con el resto de virus ADN [67,68]. La razón de este potencial y su alto ratio de sustitución es una de las grandes incógnitas de estos virus ya que, a diferencia de los virus ARN, estos virus confían su replicación a las proteínas celulares del hospedador que se caracterizan por su fiabilidad y capacidad de corrección de errores frente a otras polimerasas [69].

El aumento de la detección de virus de genoma circular y cadena simple puede estar atribuido a la incorporación de la RCA en la preparación del ADN para secuenciación. En esta técnica se emplea la polimerasa del bacteriófago  $\phi 29$  en una reacción isotérmica para amplificar genomas



circulares [70], y es tan eficiente que permite la recuperación de suficiente ADN para su secuenciación incluso a baja concentración de éste [71].

Las características genéticas de los virus CRESS-DNA son variables aunque generalmente son menores de 6 Kb y codificantes para un máximo de diez proteínas. Gracias a la secuenciación completa de los genomas circulares se ha podido determinar la arquitectura del genoma que, pese a la gran diversidad existente, presenta siempre dos proteínas esenciales para estos virus: la proteína involucrada en la replicación por círculo rodante (REP) y la que codifica para la cápsida viral (CAP).

La proteína REP es probablemente la más conservada debido a su importante función en la replicación. Esta proteína REP se caracteriza por presentar dos dominios funcionales, un dominio HUH (His-aminoácidos hidrófobos-His) organizador de la endonucleasa, situado en el extremo N-terminal y el dominio de helicasa de la superfamilia 3 (SF3) en el extremo C-terminal.

El dominio HUH se caracteriza por presentar los motivos I [FT(L/I)NN], II [PHLQG]y III [YC(S/x)K] que forman parte de la iniciación y terminación de la replicación [72,73]. El dominio de la helicasa SF3 contiene los motivos Walker A [G(P/x)(P/x)GxGK(S/t)], Walker B [uuDDD] y el motivo C [uTSN], representando “x” cualquier residuo y “u” un aminoácido hidrofóbico, estos motivos teóricamente permiten a la proteína REP actuar como helicasa durante la elongación [74]. A diferencia de la proteína REP, la CAP presenta una alta diversidad con facilidad para recombinar con otras CAP de otros virus CRESS DNA [75] y debido a su complejidad es menos conocida que la REP.

Dentro de los virus CRESS-DNA, la familia *Circoviridae* es la que agrupa a los virus de este tipo cuyo hospedador son animales. De acuerdo al criterio del ICTV esta familia está compuesta por los géneros *Circovirus* y *Cyclovirus* [76]. La mayoría de lo que conocemos sobre ellos es gracias a la investigación veterinaria, ya que hasta 2010 los cerdos eran los únicos mamíferos conocidos afectados por el género *Circovirus* [77]. Recientemente, gracias a la metagenómica, fue descrita la presencia de circovirus en otros mamíferos como murciélagos [78–81], chimpancés [82], perros [83], visones [84] y el ser humano [82,85–92].

Cada género de la familia *Circoviridae* se distingue por la organización de su genoma ambisentido. Los circovirus codifican la proteína REP en sentido positivo y la CAP en antisentido, en cambio los cyclovirus poseen la orientación contraria. Una de la estructuras más características de los

circovirus es la existencia en la región intergénica localizada en los extremos 5' de REP y CAP de una estructura en forma de lazo [30,31], en cuyo ápice se presenta un nonámero conservado [(T/n)A(G/t)TATTAC] donde se encuentra el origen de replicación.

Las características genéticas descritas anteriormente ayudan a establecer la familia a la que puede pertenecer un nuevo virus CRESS-DNA, siendo el umbral para la declaración de nueva especie un porcentaje de identidad de los nucleótidos del genoma completo inferior al 80% [93].

El potencial patogénico del género *Circovirus* en mamíferos es conocido especialmente en el caso de los virus porcinos. Uno de los más relevantes es el circovirus porcino 2 (PCV2), responsable de la epidemia del síndrome fatal en lechones [94] y agente causal de un conjunto de síndromes conocido como PCAD (enfermedad porcina asociada a circovirus), que incluye enfermedad respiratoria, dermatitis porcina, nefropatías y problemas reproductivos [95]. Este virus se transmite horizontalmente de forma respiratoria, oral, por secreciones urinarias y heces [96,97].

A pesar de todo sigue sin conocerse el rol que pueden desempeñar en el ser humano. Su relación con algunos estudios con casos graves de encefalitis, paraplejía o afección respiratoria [86,88,98,99] debe ser aún esclarecida.

Recientemente una nueva familia de virus ha sido propuesta, la familia *Redondoviridae* [99]. En estos nuevos virus se han descrito características comunes a la familia *Circoviridae* destacando una secuencia del nonámero conservado (TATTATTAT) propia de la nueva familia. A pesar de las similitudes estructurales, los análisis de la distancia por pares de la secuencia de aminoácidos de estos virus con el circovirus más parecido ("Porcine-stool-associated circular virus 5". *GenBank*: NC\_023878 ) es del 82,3 % en CAP y del 54% en REP, mostrando la posibilidad de estar ante una nueva familia dentro del grupo de los virus CRESS-DNA [99]. De acuerdo al análisis de la diversidad viral de la proteína REP de los virus de esta nueva familia se han establecido dos géneros llamados *Vientovirus* y *Brisavirus* [99]. Clínicamente han sido mayoritariamente asociados a periodontitis, y en algunos casos se han hallado en pacientes ingresados con infección respiratoria aguda en la unidad de cuidados intensivos (UCI) [99].

Los virus CRESS-DNA representan una de las oportunidades que el estudio del viroma respiratorio puede ofrecer para aumentar y profundizar en el conocimiento de la infección respiratoria aguda, incrementando nuestra capacidad de encontrar nuevos virus y de profundizar en las relaciones virus- hospedador y el estudio de la infección.

## 1.3 Aplicación de los métodos de secuenciación de nueva generación en la vigilancia de gripe

Los virus de la gripe son uno de los patógenos respiratorios más estudiados debido a su potencial epidémico y pandémico.

Estos virus pertenecen a la familia *Orthomyxoviridae* y presentan un genoma ARN de polaridad negativa de cadena simple y segmentada. Esta familia está formada por 7 géneros, de los cuales 4 de ellos engloban a diferentes tipos de virus de la gripe: el género *Alphainfluenza virus* (virus de la gripe A), el *Betainfluenzavirus* (virus de la gripe B), el *Gammainfluenzavirus* (virus de la gripe C) y el nuevo género *Deltainfluenzavirus* (virus de la gripe D) [100,101]

La variabilidad característica de estos virus se basa por un lado en la falta de capacidad correctora de su ARN polimerasa ARN dependiente, que favorece la introducción de mutaciones, y por el otro en la posibilidad de reorganización o intercambio de genes, debido a su genoma segmentado, favoreciendo la evasión del sistema inmune [102]. En los virus de la gripe se identifican, por tanto, dos de los tres mecanismos de variabilidad genética descritos en los virus ARN. Por una parte, la capacidad de introducir mutaciones (deleciones, inserciones y cambios nucleotídicos), y por otra, la reorganización o el intercambio de genes.

La variabilidad genética y antigénica puede configurar partículas virales con mayor eficacia biológica (*fitness viral*) permitiendo su circulación, mayoritariamente o no, entre una población susceptible. La vigilancia virológica de la gripe se basa en la comparación de las características genéticas y antigénicas de los virus circulantes con aquellas de los virus de referencia de los que proceden, o con los virus que utilizados para formular la vacuna estacional de la gripe de cada temporada.

Los virus de la gripe están compuestos por ocho segmentos genéticos independientes de ARN de polaridad negativa que codifican las diferentes proteínas virales: los segmentos 1, 2 y 3, que codifican la tres subunidades de la ARN polimerasa dependiente de ARN, la polimerasa básica 2 (PB2), polimerasa básica 1 (PB1) y polimerasa ácida (PA); el segmento 5 de la nucleoproteína (NP); el segmento 7 de la matriz (M), que codifica la proteína de la matriz 1 (M1) y la proteína de superficie (M2), que actúa como canal iónico; el segmento 8 de la proteína no estructural (NS), que codifica tanto la proteína 1 (NS1) relacionada con la evasión inmunitaria como la proteína

nuclear de exportación (NEP o NS2), que regula la salida de los complejos de ribonucleoproteína viral del núcleo celular; y finalmente los segmentos 4 y 6, que codifican las glicoproteínas de superficie hemaglutinina (HA) y neuraminidasa (NA), encargadas del inicio de la infección y de la liberación de partículas virales.

La identificación de los diferentes subtipos de gripe A se basa en los antígenos de superficie. En la actualidad se reconocen 18 subtipos de hemaglutinina (HA) y 11 subtipos de neuraminidasa (NA), mientras que la gripe B no tiene subtipos [103]. En el caso de gripe B dos linajes, antigénicamente y genéticamente distintos, cocirculan desde la década de los 80. Por un lado los virus que pertenecen al linaje Victoria (similares a la cepa “B/Victoria/2/87”) y por otro los que pertenecen al linaje Yamagata (similares a la cepa “B/Yamagata/16/88”) [104].

Los virus de gripe C se detectan con menos frecuencia y suelen causar infecciones leves [105], fundamentalmente en pacientes pediátricos [106], por lo que no se incluyen de forma rutinaria en los informes de la red de vigilancia de gripe.

Los virus de la gripe D, que son el último miembro que se ha incorporado a la familia *Orthomyxoviridae*, afectan principalmente al ganado y no parecen ser causa de infección ni enfermedad en el ser humano [107].

Los virus de la gripe A y B son patógenos humanos asociados con hospitalización y mortalidad. En España durante la temporada 2017/2018 circularon de forma conjunta gripe A (H3N2) y gripe B produciendo 52000 hospitalizaciones con 3000 admisiones en la unidad de cuidados intensivos (UCI) y cerca de 15000 muertes relacionadas con el virus [108]. Estos datos fueron obtenidos por la red de vigilancia de gripe cuyo objetivo es relacionar las características epidemiológicas, clínicas y virológicas de la infección gripal en la población, así como la obtención de virus circulantes para su análisis antigénico como posibles candidatos para formar parte de la vacuna de la temporada siguiente. Debido a su importancia en salud pública, la caracterización virológica es crítica para conocer las propiedades genéticas y antigénicas de los virus circulantes [6].

A través del sistema de vigilancia son analizados los virus circulantes, y sus secuencias comparadas con las de los genes HA y NA de los virus incluidos en la vacuna de la temporada, evaluándose así las diferencias que pueden aparecer por la deriva genética de los virus gripales. Durante el transcurso de una o varias temporadas epidémicas, los virus circulantes presentarán

cambios genéticos con respecto a los virus incluidos en la vacuna correspondiente (deriva genética). A veces, estos cambios genéticos conllevan cambios antigénicos que hacen que los anticuerpos que se han generado frente a los virus vacunales no reconozcan a los virus circulantes. Actualmente, la estimación de la distancia antigénica basada en la secuencia de una determinada proteína y en las diferencias observadas en los análisis de inhibición de la hemaglutinación siendo ambos sistemas de enorme utilidad para la investigación de los virus gripales, y como las bases de datos contienen cientos de miles de secuencias de virus gripales la predicción de la discordancia antigénica da lugar a estudios que modelizan estos aspectos y los aplican a la predicción de la efectividad vacunal y de cómo los anticuerpos producidos ante una primoinfección gripal condicionan la respuesta del individuo por las reacciones cruzadas de sus anticuerpos [109].

La secuenciación de genomas es un proceso que determina el orden o la secuencia de los nucleótidos (A, U/T, C o G) en cada uno de los genes presentes en el genoma de un virus. En el caso de los virus de la gripe, la secuenciación completa de sus ocho segmentos resulta en una secuencia total de alrededor de 13.500 nucleótidos, mientras que la secuenciación parcial muestra la secuencia de algunos segmentos o genes, o incluso determinadas partes de ellos que son seleccionadas por su concordancia con las características genéticas del gen completo para su estudio. Este sistema de análisis se denomina “caracterización genética” y se realiza en los laboratorios de salud pública de todo el mundo desde los años 1980. Las secuencias genéticas obtenidas se almacenan en bases de datos públicas, como *GenBank* (<https://www.ncbi.nlm.nih.gov/genbank/>), y desde hace unos años también en la “*Global Initiative on Sharing All Influenza Data*” (*GISAID*; <https://www.gisaid.org/>). Las secuencias almacenadas permiten comparar los genes completos, o fragmentos de ellos, de los virus de la gripe circulantes durante la temporada epidémica con las secuencias de los genes de los virus gripales que circularon anteriormente y con las secuencias de los virus componentes de las vacunas (virus vacunales o referencia). Mediante estas comparaciones de secuencias genéticas se obtienen conclusiones sobre: 1. cómo los virus se relacionan entre sí; 2. cómo los virus evolucionan a lo largo del tiempo; 3. cuáles son las variaciones genéticas o mutaciones que aparecen cuando los virus se diseminan ocasionando enfermedades más graves o desarrollando resistencia a los medicamentos antivirales; 4. qué grado de semejanza presenta un determinado virus de la gripe con el incluido como componente de la vacuna estacional de la gripe y 5. cuáles

son las adaptaciones que se producen en los virus animales que podrían causar un salto de la barrera interespecie e infectar al ser humano.

Actualmente la secuenciación del genoma del virus de la gripe se realiza mediante dos técnicas en los laboratorios de salud pública: la secuenciación por Sanger [110] y las técnicas de secuenciación de nueva generación [111–114] .

El método de secuenciación Sanger es el método que se utiliza tradicionalmente para conocer las características genéticas y la evolución de los virus de la gripe como parte de la vigilancia virológica. Cuando la secuenciación de Sanger es aplicada en un aislamiento viral se identifica la secuencia genética que predomina y que ha sido seleccionada entre todas las posibles variantes del virus. La aplicación de la secuenciación de Sanger directamente en muestras clínicas permite obtener la secuencia de los virus presentes en la muestra. Sin embargo, las variantes que estén presentes en baja proporción en una población viral no aparecerán representadas en la secuencia final. En muchos laboratorios de salud pública sólo se analiza la subunidad HA1 del segmento que codifica la HA y se obtienen secuencias parciales de ese segmento que son lo suficiente informativas como para caracterizar los virus.

Desafortunadamente, este método a pesar de su fiabilidad es costoso, ya que el diseño de los oligonucleótidos cebadores para la amplificación previa a la secuenciación suele estar basado en pequeñas regiones relativamente conservadas de los diferentes segmentos genómicos, por lo que se requiere el diseño de varios conjuntos de oligonucleótidos y de PCRs específicas para cada uno de los segmentos virales [110]. La estrategia del diseño de cebadores específicos puede ser inefectiva ya que dificulta la detección de nuevos subgrupos genéticos (clados) de gripe o incluso variantes de un mismo subgrupo presentes en la muestra debido a la posibilidad de una falta de ajuste en la secuencia por parte de los oligonucleótidos cebadores.

Probablemente la diferencia más importante de esta técnica con respecto a las técnicas NGS es que solo permite la detección de la secuencia consenso del virus dominante en un momento determinado obviando la diversidad de subpoblaciones y la dinámica de transmisión del virus [115,116].

El uso de técnicas NGS ha revolucionado el estudio de los virus de la gripe al reforzar la habilidad de los laboratorios de salud pública para: 1. conocer las características genómicas de los virus circulantes y su diversidad; 2. la posible emergencia de virus gripales genéticamente diferentes

con potencial pandémico; 3. detectar cambios genéticos que conlleven resistencia a antivirales, y 4. optimizar la selección de cepas para ser elegidas en la composición de las vacunas [20,117].

La existencia de oligonucleótidos que permiten la secuenciación de genomas completos que permiten a su vez la representación de variantes minoritarias como los protocolos diseñados para amplificar el virus completo para gripe A [16], gripe B [17] o ambos tipos con un solo set de oligonucleótidos [18] ha supuesto un gran avance en la secuenciación de genoma completo del virus de la gripe para su secuenciación NGS. Al desarrollo y mejora de técnicas de preparación de muestras y la aparición de PCRs para la amplificación de genoma completo, se une la rápida evolución de la tecnología de secuenciación que ofrece varias plataformas de análisis bioinformático con diferente potencial tal y como hemos mencionado en el apartado de metagenómica clínica.

Recientemente la disponibilidad de dos plataformas de análisis bioinformático desarrolladas con estrategias diferentes como son *IRMA* [21] e *INSaFLU* [25], específicas para gripe y asequibles para el usuario, ha facilitado el análisis de las secuencias en laboratorios de referencia relacionados con salud pública. La plataforma *INSaFLU* está basada en el mapeo de lecturas contra una secuencia de referencia especificada por el usuario para generar una secuencia consenso [25]. En cambio, *IRMA* es un meta-ensamblador que va cambiando la secuencia molde según se vaya pareciendo a la lectura de la secuencia (refinamiento iterativo) para incrementar la sensibilidad y ensamblar la secuencia consenso sin necesidad de aportar una secuencia de referencia [21]. Se pueden utilizar otras plataformas de análisis informático online como *Virus TAP* [24] que tiene en común con *INSaFLU* el método de generación de secuencias consenso por ensamblaje *de novo*, aunque no es específica para gripe.

En este trabajo se han desarrollado cada uno de los pasos necesarios para implementar de manera exitosa la secuenciación masiva aplicada a un laboratorio de referencia de gripe seleccionando cada una de las técnicas desde el enriquecimiento hasta el análisis bioinformático.

La implementación de estas técnicas no son únicamente importantes para mejorar el diagnóstico e investigación, sino que forman parte de la respuesta del laboratorio de Virus Respiratorios y Gripe a la estrategia del Centro Europeo para el Control y Prevención de Enfermedades (ECDC, según sus siglas en inglés) 2019-2021 [118]. El ECDC considera que las técnicas de secuenciación de genoma completo (WGS, según sus siglas en inglés) tienen potencial para mejorar el tipado

molecular y la epidemiología de gripe desarrollando la investigación de brotes y epidemias, el estudio de dinámicas de transmisión y la predicción de resistencias antivirales o de virulencia del patógeno. Por ello, en el nuevo marco del ECDC para la integración de estas técnicas genómicas en vigilancia ha sido incluido el virus de la gripe.

## 1.4 Secuenciación de adenovirus por métodos de secuenciación de nueva generación

El desarrollo de los métodos NGS para muestras respiratorias permite el estudio de viroma en muestras procedentes de cualquier especie. Gracias a la disponibilidad de muestras respiratorias de murciélagos permitió ampliar los objetivos de esta tesis.

Los murciélagos son el segundo orden más numeroso de mamíferos, agrupando a más de 1200 especies diferentes [119]. Su estudio en términos virológicos resulta interesante ya que su gran ubicuidad y su organización en grupo predisponen a la infección y diseminación viral [120]. El orden Chiroptera es un reservorio conocido de varias familias de virus como las familias *Rhabdoviridae*, *Coronaviridae*, *Herpesviridae*, *Filoviridae*, *Reoviridae*, *Paramyxoviridae* y *Astroviridae* entre otras [121].

En España la vigilancia de rabia es un asunto de salud pública importante debido a la situación geográfica entre África y Europa en la que se encuentra la Península Ibérica [122]. Estudios sobre la presencia del virus de la rabia en murciélagos ibéricos describieron a los murciélagos del género *Eptesicus* como vectores del virus [123,124], siendo conocido el flujo genético de murciélagos *Eptesicus isabellinus* entre el sur de España y el norte de Marruecos [125], el conocimiento y control virológico de los virus de murciélago es una cuestión principal de salud pública en España.

Debido al material biológico disponible gracias al sistema de vigilancia del virus de la rabia se realizaron estudios en la península que permitieron la descripción de la presencia de nuevos virus como el lisavirus Lleida [126], el filovirus Lloviu detectado en *Miniopterus schreibersii* muertos en el Norte de España [127], 14 nuevos coronavirus incluyendo 2 betacoronavirus relacionados con el grupo del coronavirus MERS [128], 42 potenciales nuevos betaherpesvirus y



10 potenciales nuevos rhabdovirus procedentes de las familias Vespertilionidae, Miniopteridae, Rhinolophidae, Molosidae and Pteropodidae en el sur y norte de España [129,130].

Dentro de los virus respiratorios humanos cuyo reservorio conocido son los murciélagos se encuentra la familia *Adenoviridae* [131–134]. Los adenovirus se subdividen en 5 géneros: *Mastadenovirus*, *Aviadenovirus*, *Atadenovirus*, *Siadenovirus* y los *Ictadenovirus* [135].

La detección de AdV en murciélagos comienza en 2008 con el aislamiento del primer adenovirus de murciélago, BtAdV1-FBV1, durante el intento de establecer una línea celular de un murciélago *Pteropus dasymallus yayeyamae* en Japón [136].

A pesar de la detección del BtAdV1-FBV1 hasta 2009 no fue caracterizado un AdV aislado en muestras de murciélagos del suborden Microchiroptera, el único presente en Europa. El análisis de 55 murciélagos en libertad de la familia *Vespertilionidae* en Alemania caracterizó el segundo AdV, el “BtAdV-2 cepa PPV1”, fue identificado en tres murciélagos *Pipistrellus pipistrellus* [137], siendo el primer AdV descrito en un microquiróptero pero el segundo completamente secuenciado [138].

El primer AdV completamente secuenciado procedente de un microquiróptero fue el “BtAdV-3 cepa TJM” encontrado en las muestras obtenidas de un murciélago *Myotis ricketti* [139]. Estos nuevos AdVs fueron renombrados a propuesta del ICTV como *mastadenovirus A de murciélago* (BtAdV-3 cepa TJM) y *mastadenovirus B de murciélago* (BtAdV-2 cepa PPV1).

Varios estudios han mostrado una gran diversidad viral mundial siendo detectados AdVs en murciélagos analizados en Brasil [140], Japón [136], Alemania [131,131,138,141], China [139,142], Hungría [133,141], Ghana [143], Zambia [144], Kenia [134], Sudáfrica [145] y EEUU [146]. Como consecuencia de estos estudios han sido propuestas al ICTV nuevas especies de adenovirus de murciélago como el *mastadenovirus C* (*Rhinolophus sinicus* WIV9, KT698853) [133], el *mastadenovirus D* (*Miniopterus schreibersi* WIV12, KT698856) [142], el *mastadenovirus E* (*M. schreibersi* WIV13, KT698852) [142], *mastadenovirus F* (*Rousettus leschenaultii* WIV17, KX961095) [142] y el *mastadenovirus G de murciélago* (*Corynorhinus rafinesquii* 250-A, KX871230) [146].

Estos trabajos han incrementado el conocimiento sobre nuevos virus y su potencial como patógenos humanos. Debido al papel activo de los murciélagos como reservorios virales la

caracterización y conocimiento de estos virus es una parte importante de la vigilancia de salud pública.

El estudio de AdVs de murciélagos tiene el potencial de no solo aportar información valiosa para la vigilancia de salud pública sino que la secuenciación del genoma completo por NGS permite la caracterización de los nuevos mastadenovirus [138,139,142,146], el desarrollo de protocolos de trabajo para la secuenciación completa de virus ADN de gran tamaño y la posibilidad conocer nuevos aspectos de la infección de adenovirus.

El análisis por metagenómica de virus ADN de gran tamaño no es sencillo y se complica aún más en el caso de la muestra respiratoria debido a la baja carga viral de las muestras clínicas que dificulta su detección y la posibilidad de obtención del genoma completo.

Trabajos previos han logrado secuenciar AdVs completos usando técnicas NGS tanto en virus de murciélagos como en virus humanos [138,139,142,146–148]. Estos trabajos tienen un factor común: o parten de cultivo celular o de órganos internos de animales enfermos, en ambos casos presentándose muestras con gran cantidad de virus disponible para secuenciar.

El cultivo celular es una técnica sensible cuyos resultados dependen de varios factores como la carga viral, la calidad de la muestra, la línea celular elegida o la experiencia en el manejo de la técnica [149]. Un trabajo previo estimó que únicamente el 2% de las muestras respiratorias pediátricas negativas para PCRs de diagnóstico respiratorio fueron recuperadas tras el cultivo celular [150], su bajo rendimiento junto con la duración de la técnica hace necesario el desarrollo de protocolos para poder secuenciar genoma completo directamente de muestra clínica.

Como hemos mencionado previamente el enriquecimiento de virus es uno de los métodos empleados para aumentar la proporción de secuencias virales y en el caso de AdV es la mejor opción ya que a diferencia de otros virus como la gripe no se dispone de un conjunto de cebadores para la amplificación del genoma completo.

El enriquecimiento viral basado en la centrifugación, la filtración y el tratamiento enzimático con el fin de eliminar partículas de mayor tamaño que las virales y el ADN y ARN del hospedador es un método sencillo pero a pesar de ello su rendimiento en muchos casos es menor que el obtenido por otras técnicas de enriquecimiento como los métodos de captura por sondas.

Los métodos de captura por sondas han sido ya empleados para el enriquecimiento de diversos virus respiratorios [151,152]. Dentro del desarrollo de estas técnicas han sido empleados tanto paneles de captura caseros y específicos para un conjunto de virus, como por ejemplo un panel de virus respiratorios [152,153], como el panel comercial y de amplio espectro conocido como panel mediante captura por sondas *HyperCap Target Enrichment* [151].

La obtención de genoma completo es imprescindible para proponer nuevas especies de AdV al ICTV pero además es importante para conocer otros aspectos de la infección por AdV, ya que para muchos serotipos humanos no ha sido identificado el receptor celular involucrado en la infección y dado que estos virus replican pobremente en animales la comprensión de la infección natural puede incrementar el conocimiento de las infecciones de AdV [154] .

Los AdVs humanos causan un amplio rango de síndromes clínicos y han sido diagnosticados en casos de neumonía fatal, cistitis hemorrágica o enfermedad diseminada en trasplantados de médula [155–157], por lo que el estudio de virus completo, con el fin de conocer aspectos aún desconocidos de estos virus, puede tener un impacto importante en pacientes tan vulnerables como los pacientes trasplantados o los pacientes pediátricos.

# Hipótesis y Objetivos



## HIPÓTESIS Y OBJETIVOS

El desarrollo y adecuación de las técnicas moleculares basadas en los métodos de secuenciación de nueva generación aplicados al análisis directo de muestras clínicas respiratorias, tanto humanas como animales, es de enorme utilidad a la hora de investigar las características genómicas de los virus respiratorios detectados en dichas muestras como componentes del viroma respiratorio. La capacitación tecnológica del Laboratorio de Referencia Nacional de Gripe y Virus Respiratorios permite abordar el estudio de la infección respiratoria viral de manera integral. La aplicación de estas técnicas de secuenciación de nueva generación aporta resultados que mejoran la caracterización del viroma respiratorio. Existen grandes retos en el conocimiento de ecosistemas tan complejos como es el tracto respiratorio y abordando diferentes situaciones epidemiológicas prácticas se puede conseguir una visión de las posibilidades tecnológicas y sus avances en diferentes campos como es el caso del estudio de viroma de pacientes pediátricos, la detección y el descubrimiento de nuevos virus respiratorios, la mejora del conocimiento genómico de los virus de la gripe o el estudio de la presencia de virus respiratorios en hospedadores diferentes al ser humano.

### Objetivo general:

Desarrollar e implementar las técnicas de secuenciación de nueva generación en el estudio de los virus respiratorios presentes en las muestras respiratorias para conocer la composición viral del tracto respiratorio y sus relaciones ecológicas.

### Objetivos concretos:

**Objetivo I: Desarrollo de un protocolo para el estudio de viroma en las muestras clínicas respiratorias.**

1.1 Ajuste y desarrollo de los diferentes protocolos de enriquecimiento y de extracción de ácidos nucleicos.

1.2 Aplicación de la amplificación independiente de secuencia con un cebador único para el estudio del viroma respiratorio.

**Objetivo II: Aplicación de los métodos desarrollados en el estudio de la composición del viroma respiratorio y las relaciones ecológicas de sus integrantes**

2.1 Análisis del viroma de muestras de seguimiento de pacientes pediátricos

2.2 Estudio de la asociación entre virus detectados.

2.3 Estudio del viroma respiratorio como herramienta para descubrir posibles nuevos virus respiratorios.

**Objetivo III: Desarrollo y aplicación de técnicas de secuenciación de nueva generación en el estudio y referencia de los virus gripales y su vigilancia epidemiológica.**

3.1 Desarrollo de un sistema de fácil implementación en el Laboratorio de Referencia Nacional de Gripe.

3.2 Adecuación de protocolos recomendados por OMS y CDC y sus análisis bioinformáticos.

**Objetivo IV: Estudio de virus respiratorios ADN de gran tamaño y complejidad genómica detectados en hospedadores no humanos: nuevos adenovirus en murciélagos ibéricos.**

4.1 Conocer las relaciones filogenéticas y evolutivas entre los adenovirus de diferentes especies de murciélagos de la Península Ibérica

4.2 Secuenciar los nuevos adenovirus usando técnicas de secuenciación masiva tras comparación de diferentes métodos de enriquecimiento y preparación de librerías

# Materials





## MATERIALES

### 3.1. Líneas celulares

- Células de riñón de mono verde (VERO C1008 [Vero 76, clone E6, Vero E6]): adquiridas en la colección de cultivos tipo americana (ATCC). Referencia ATCC® CRL-1586™. Fueron usadas para el aislamiento del virus respiratorio circular humano, cuyas siglas en inglés son HRCiV, y adenovirus de murciélago.
- Fibroblastos de pulmón de embrión humano (L132): adquiridas en la ATCC. Referencia: ATCC® CCL5™. Fueron usadas para el aislamiento del virus HRCiV.
- Células humanas Hep-2: Adquiridas en la ATCC. Referencia ATCC® CCL-23™. Fueron empleadas en el aislamiento del virus HRCiV y adenovirus de murciélago
- Células de riñón de perro con expresión aumentada de ácido siálico (MDCK-SIAT1): Adquiridas en la Colección Europea de Cultivos Celulares Autenticados (ECACC). Referencia: ECACC 05071502. Usadas para el aislamiento del virus de la gripe y de HRCiV.

### 3.2. Materiales de referencia utilizados como control virológico

Los materiales de referencia de los virus pertenecientes a las familias *Adenoviridae*, *Orthomyxoviridae*, *Circoviridae*, *Paramyxoviridae*, *Parvoviridae*, *Picornaviridae* y *Coronaviridae* fueron utilizados como controles positivos en las técnicas de PCR descritas en la Tabla 1 para la detección e identificación de diferentes genes virales.

Familia	Virus	Origen del virus	Técnica utilizada
<b><i>Adenoviridae</i></b>	AdV 41	Cepa de referencia WS-91 (ATCC)	PCRs en los genes de la polimerasa y hexón
<b><i>Orthomyxoviridae</i></b>	Gripe A (H1N1pdm)	Cepa de referencia A/California/07/2009 (OMS)	PCR Múltiple tiempo real Gripe ABC PCR multisegmento A PCR multisegmento universal
	Gripe B (linaje Yamagata)	Cepa de referencia B/Phuket/3073/2013 (OMS)	PCR Múltiple tiempo real Gripe ABC PCR multisegmento B PCR multisegmento universal

<b><i>Circoviridae</i></b>	HRCiV	Muestra clínica respiratoria	PCR de estudio HRCiV 2 PCR para genoma completo de HRCiV
<b><i>Paramyxoviridae</i></b>	VRSB	Cepa de referencia VRSB-CH (ATCC)	PCR Múltiple tiempo real BRQ
	HMNV	Muestra clínica respiratoria	
<b><i>Parvoviridae</i></b>	BOV	Muestra clínica respiratoria	PCR Múltiple tiempo real BRQ
<b><i>Picornaviridae</i></b>	EV	Cepa de referencia Echovirus 30 Bastiani (ATCC)	PCR Múltiple tiempo real Virus Respiratorios 2
	RIN	Cepa del ATCC de Rinovirus 2 (ATCC)	
<b><i>Coronaviridae</i></b>	CoV 229E	Aislado de muestra clínica	PCR genérica coronavirus
	CoV MERS	Plásmidos de ADN (OMS)	PCRs tiempo real en gen UpE y en gen ORF1A de Coronavirus MERS

**Tabla 1:** Relación de virus control empleados en las diferentes técnicas utilizadas

### 3.3 Sujetos de estudio y muestras biológicas

Las muestras analizadas en los diferentes objetivos de este trabajo procedían fundamentalmente de dos vías: las muestras obtenidas para el diagnóstico de gripe y las muestras obtenidas de pacientes pediátricos durante el desarrollo de los proyectos de investigación.

Las muestras recibidas en el Laboratorio de Virus Respiratorios y Gripe para el diagnóstico de gripe procedían de hospitales del sistema nacional de salud de España, así como de la Red de Médicos Centinela del Sistema de Vigilancia de la Gripe en España.

Las muestras pediátricas con las que se realizó el estudio metagenómico fueron obtenidas de pacientes atendidos en los Servicios de Pediatría del Hospital Universitario Severo Ochoa de Leganés y del Hospital Universitario La Paz de Madrid, durante el desarrollo de los proyectos de investigación “Metagenómica viral asociada al estudio de la infección grave respiratoria” (PI15CIII/00028) y “Los virus respiratorios en las infecciones respiratorias agudas de recién nacidos y pacientes pediátricos: asociación con la gravedad del episodio agudo y con el desarrollo de asma y sibilancias recurrentes a largo plazo.” (PI18CIII/00009).

Las diferentes tipos de muestra recibidas en el Laboratorio de Virus Respiratorios y Gripe y empleadas en los diferentes objetivos de este trabajo son descritas en la Tabla 2.

Tipo de muestra	Código
Aspirado bronquial	AB
Aspirado brocoalveolar	ABA
Aspirado nasofaríngeo	ANF
Aspirado traqueal	AT
Aspirado traqueobronquial	ATB
Espuito	Espuito
Exudado nasofaríngeo	ENF
Exudado faríngeo	EF
Lavado nasal	LN
Lavado nasofaríngeo	LNF
Lavado broncoalveolar	LBA
Tejido pulmonar (biopsia o necropsia)	Biopsia/Necropsia
Cepa/aislado de virus	Cepa
Líquido alantoideo	Alanto
Líquido pleural	LQ pleural
Líquido cefalorraquídeo	LCR
Heces	Heces
Orina	Orina
Medula ósea	MO
Exudado conjuntival	EC
Humor vítreo	Humor vítreo
Humor acuoso	Humor acuoso
Líquido pericárdico	LQ PER
Líquido amniótico	LQ AMN
Líquido sinovial	LQ SIN

**Tabla 2:** Tipo de muestras recibidas en el Laboratorio de Virus Respiratorios y el código de identificación de dichas muestras

### 3.3.1 Características de las muestras empleadas para el desarrollo de un protocolo para el estudio de viroma en las muestras clínicas respiratorias

Para la estandarización inicial del protocolo de viroma fueron empleados 10 ANFs procedentes de pacientes pediátricos. Previamente a la estandarización del protocolo para el análisis de viroma, los ANFs usados fueron analizados en el Laboratorio de Virus Respiratorios y Gripe empleando las PCRs de detección de virus respiratorios desarrolladas en el laboratorio: PCR múltiple tiempo real Gripe ABC [158], PCR múltiple tiempo real BRQ [158] y PCR múltiple tiempo real Respiratorios 2 [159]. El empleo de estas PCRs permitió conocer la presencia o ausencia y el tipo o tipos de virus respiratorios presentes en la muestra.

Las muestras respiratorias empleadas en el objetivo I fueron identificadas como: SO14472, SO14766, SO14513, SO14632, SO14381, SO14278, SO14389, SO14509, SO14713 y SO14842. El

diagnóstico viral obtenido por las PCRs en tiempo real de virus respiratorios descritas mostró que 2 de las muestras eran positivas para virus ADN y 8 para virus ARN (Tabla 3).

Identificación de muestra	Diagnóstico viral
SO14472	RV
SO14766	AdV
SO14513	BoV
SO14632	VRS
SO14381	VRS
SO14278	RV
SO14389	RV
SO14509	RV
SO14713	RV
SO14842	RV

**Tabla 3:** Identificación y caracterización virológica de virus respiratorios obtenidos por métodos moleculares en las muestras empleadas en el objetivo I

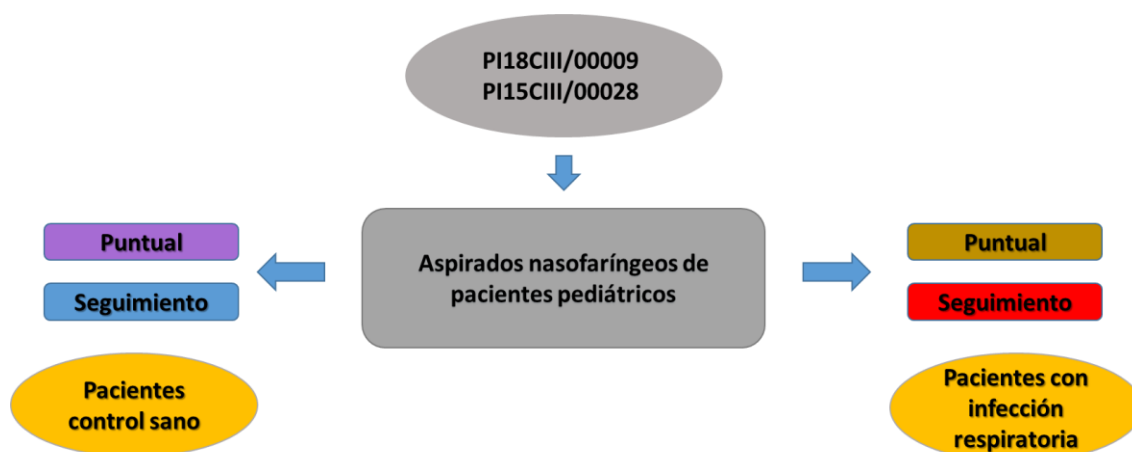
### 3.3.2 Características de las muestras empleadas en el estudio de la aplicación de los métodos desarrollados en el estudio de la composición del viroma respiratorio y las relaciones ecológicas de sus integrantes

Para el estudio de la composición del viroma respiratorio fueron estudiados 153 ANFs procedentes de pacientes pediátricos.

Las muestras respiratorias pediátricas empleadas se dividieron en dos grandes grupos (Figura 2): muestras tomadas de forma puntual a un paciente o muestras de seguimiento del mismo paciente a lo largo del tiempo. Tanto las muestras de toma puntual como las muestras con seguimiento procedían de pacientes con infección respiratoria y pacientes del grupo control sano.

El grupo de los individuos control sanos fueron niños sin sintomatología respiratoria que acudían al hospital por causas ajenas a la infección respiratoria, en el caso de las muestras de seguimiento estos pacientes acudían periódicamente al hospital debido a dolencias no relacionadas con patógenos respiratorios. Las muestras de estos controles eran tomadas únicamente en el caso de que no hubiera habido ningún síntoma respiratorio diez días antes de la visita al hospital y además se confirmaba la ausencia de virus mediante las PCRs de detección de virus respiratorios: PCR múltiple tiempo real Gripe ABC [158], PCR múltiple tiempo real BRQ [158] y PCR múltiple tiempo real Respiratorios 2 [159]. Estos individuos sanos son comparables al grupo de pacientes

con IRA ya que los individuos del grupo de control se emparejaron por edad, provenían de la misma área geográfica y sus muestras también se tomaron sistemáticamente durante la misma ventana de tiempo.



**Figura 2:** Esquema de los distintos tipos muestra respiratoria empleada procedentes de los pacientes pediátricos para el estudio de viroma. Los colores se asocian con la descripción clínica en la Tabla 4

Los pacientes con infección respiratoria también fueron analizados con la batería de PCRs de detección de virus respiratorias del laboratorio descritas previamente.

La Tabla 4 describe la situación clínica y la edad de los pacientes estudiados clasificando por colores las 57 muestras puntuales de pacientes con infección respiratoria (amarillo), las 40 muestras puntuales de individuos control sano (morado), las 35 muestras de seguimiento de pacientes con infección respiratoria (rojo) y las 21 muestras de seguimiento de pacientes control (azul).

Muestra	Edad	Fiebre	Hipoxia	Per. Hosp.	Rx. Tórax	Diagnóstico	Antibióticos	UCI
1	53,3	No	No	2	N/A	Bronquiolitis	No	No
2	55,8	No	Sí	4	Sin alteraciones	Asma	No	No
3	55,8	Sí	Sí	3	Sin alteraciones	Bronquiolitis	No	No
4	41,4	Sí	No	3	Infiltrados	Neumonía	Sí	No
5	41,4	No	Sí	2	N/A	Asma	No	No
6	34,6	Sí	Sí	4	Infiltrados	Asma	No	No
7	34,3	No	No	2	N/A	ITRS	No	No
8	34,3	No	No	N/A	N/A	ITRS	No	No
9	23,6	No	No	3	Sin alteraciones	Bronquiolitis	No	No
10	31,3	Sí	No	3	Sin alteraciones	Sibilancias	No	No
11	42,8	Sí	Sí	5	Sin alteraciones	Bronquiolitis	No	No

12	46,3	Sí	Sí	N/A	Infiltrados	Asma	No	No
13	46,3	Sí	Sí	8	Infiltrados	Asma	No	No
14	3,0	Sí	Sí	9	Sin alteraciones	Bronquiolitis	No	No
15	2,3	No	Sí	3	Infiltrados/	Sibilancias	No	No
16	2,1	Sí	Sí	2	Sin alteraciones	Bronquiolitis	No	Sí
17	2,1	Sí	Sí	8	Infiltrados	Bronquiolitis	No	No
18	50,8	Sí	Sí	5	Infiltrados	Sibilancias	Sí	Sí
19	50,8	Sí	Sí	10	Infiltrados	Bronquiolitis	No	No
20	28,1	No	Sí	7	Sin alteraciones	Sibilancias	No	No
21	29,8	Sí	Sí	3	N/A	Bronquiolitis	No	Sí
22	33,9	Sí	No	2	Infiltrados	Sibilancias	Sí	No
23	36,3	No	No	2	N/A	Sibilancias	No	No
24	42,8	Sí	Sí	3	Infiltrados	Neumonía	Sí	No
25	6,2	No	Sí	5	N/A	Sibilancias	No	No
26	7,3	Sí	Sí	4	Sin alteraciones	Sibilancias	No	No
27	21,3	No	Sí	6	Sin alteraciones	Bronquiolitis	No	No
28	6,4	No	Sí	3	N/A	Asma	No	No
29	5,4	Sí	Sí	3	Infiltrados	Sibilancias	No	No
30	5,6	No	Sí	4	N/A	Bronquiolitis	No	No
31	5,3	Sí	Sí	3	Sin alteraciones	Bronquiolitis	Sí	Sí
32	1,2	Sí	No	1	Infiltrados	Bronquiolitis	Sí	No
33	8,1	No	Sí	4	Sin alteraciones	Sibilancias	No	No
34	6,3	Sí	Sí	4	Infiltrados	Sibilancias	No	No
35	151,2	Sí	No	9	N/A	Fiebre sin foco	Sí	No
36	24,7	Sí	No	3	N/A	ITRS	Sí	Sí
37	41,4	Sí	Sí	4	Infiltrados	Neumonía	Sí	Sí
38	46,7	Sí	Sí	2	Infiltrados	Asma	No	No
39	13,6	Sí	No	2	N/A	ITRS	No	No
40	9,8	No	Sí	3	Sin alteraciones	Sibilancias	No	No
41	11,6	Sí	No	2	Infiltrados	Neumonía	Sí	No
42	13,3	Sí	No	5	Infiltrados	Neumonía	Sí	No
43	15,1	Sí	No	7	Sin alteraciones	Fiebre sin foco	No	No
44	19,2	No	N/A	5	N/A	EAM	No	No
45	20,9	No	No	2	N/A	ITRS	No	No
46	15,6	Sí	No	7	Infiltrados	Neumonía	No	No
47	34,8	Sí	No	3	Infiltrados	Asma	Sí	No
48	34,1	No	No	N/A	N/A	Bronquiolitis	No	No
49	35,8	Sí	Sí	2	Sin alteraciones	Sibilancias	No	No
50	13,7	No	Sí	2	Sin alteraciones	Asma	No	No
51	42,3	Sí	Sí	3	Infiltrados	Sibilancias	No	No
52	43,5	Sí	Sí	3	Infiltrados	Bronquiolitis	Sí	No

53	45,3	Sí	Sí	N/A	Infiltrados	Asma	No	No
54	48,8	No	Sí	2	N/A	Asma	No	No
55	53,5	Sí	Sí	5	Infiltrados	Sibilancias	No	No
56	10,7	Sí	Sí	3	Infiltrados	Bronquiolitis	Sí	No
57	3,8	No	Sí	3	Infiltrados	Asma	No	No
58	0,9	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
59	1,8	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
60	1,8	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
61	1,2	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
62	1,2	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
63	1,2	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
64	1,2	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
65	1,2	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
66	1,6	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
67	1,6	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
68	1,6	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
69	0,9	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
70	0,9	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
71	2,1	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
72	1,5	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
73	1,5	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
74	1,5	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
75	1,5	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
76	1,5	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
77	1,5	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
78	0,8	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
79	0,9	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
80	1,8	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
81	1,8	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
82	1,1	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
83	2,1	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
84	2,1	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
85	2,1	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
86	2,1	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
87	2,1	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
88	31,6	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
89	32,6	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
90	1,0	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
91	1,0	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
92	1,2	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
93	0,5	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
94	0,7	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
95	0,7	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
96	1,0	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
97	1,0	N/A	N/A	N/A	N/A	Control sano	N/A	N/A



98	N/A	Sí	No	N/A	N/A	ITRS	N/A	N/A
99	N/A	Sí	No	N/A	N/A	ITRS	N/A	N/A
100	N/A	Sí	No	N/A	N/A	ITRS	N/A	N/A
101	N/A	N/A		2	N/A	N/A	N/A	N/A
102	N/A	Sí			N/A	Sibilancias	N/A	N/A
103	N/A	No	No	N/A	N/A	Sibilancias	N/A	N/A
104	2,3	No	Sí	4	N/A	Sibilancias	N/A	N/A
105	8,0	No	Sí	2	N/A	Sibilancias	N/A	N/A
106	11,9	No	Sí	1	N/A	Sibilancias	N/A	N/A
107	6,5	Sí	Sí	N/A	N/A	ITRS	N/A	N/A
108	10,7	No	No	N/A	N/A	N/A	N/A	N/A
109	14,8	Sí	No	N/A	N/A	ITRS	N/A	N/A
110	23,7	Sí	Sí	3	N/A	ITRS	N/A	N/A
111	N/A	N/A	No	N/A	N/A	Sibilancias	N/A	N/A
112	N/A	N/A	No	N/A	N/A	ITRS	N/A	N/A
113	N/A	No		1	N/A	ITRS	N/A	N/A
114	N/A	No	No	N/A	N/A	Bronquiolitis	N/A	N/A
115	N/A	No	No	N/A	N/A	Bronquiolitis	N/A	N/A
116	N/A	No	No	N/A	N/A	Sibilancias	N/A	N/A
117	N/A	No	No	N/A	N/A	Sibilancias	N/A	N/A
118	3,8	No	No	N/A	N/A	Sibilancias	N/A	N/A
119	5,0	No	No	N/A	N/A	ITRS	N/A	N/A
120	7,6	N/A	No	N/A	N/A	Bronquiolitis	N/A	N/A
121	37,9	Sí	Sí	4	N/A	Sibilancias	N/A	N/A
122	37,9	No	No	N/A	N/A	ITRS	N/A	N/A
123	39,9	No	Sí	7	N/A	Sibilancias	N/A	N/A
124	48,6	No	No	N/A	N/A	Sibilancias	N/A	N/A
125	11,8	Sí	No	N/A	N/A	ITRS	N/A	N/A
126	18,8	No	No	N/A	N/A	ITRS	N/A	N/A
127	25,0	Si	No	N/A	N/A	Bronquiolitis	N/A	N/A
128	17,0	Sí	Sí	7	N/A	Bronquiolitis	N/A	N/A
129	30,3	No	No	N/A	N/A	ITRS	N/A	N/A
130	25,0	No	No	N/A	N/A	Bronquiolitis	N/A	N/A
131	26,3	Sí	No	N/A	N/A	Laringitis	N/A	N/A
132	48,4	No	No	N/A	N/A	Laringitis	N/A	N/A
133	2,3	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
134	4,0	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
135	5,2	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
136	6,9	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
137	0,9	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
138	2,1	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
139	2,7	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
140	4,4	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
141	1,1	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
142	2,8	N/A	N/A	N/A	N/A	Control sano	N/A	N/A

143	1,4	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
144	0,6	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
145	2,6	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
146	3,2	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
147	4,3	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
148	1,1	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
149	1,6	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
150	2,3	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
151	2,8	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
152	3,3	N/A	N/A	N/A	N/A	Control sano	N/A	N/A
153	5,1	N/A	N/A	N/A	N/A	Control sano	N/A	N/A

**Tabla 4:** Presentación clínica de los 153 pacientes pediátricos empleados en el estudio del viroma respiratorio. La edad está expresada en meses y el período de hospitalización está expresado en días

### 3.3.3 Características de las muestras clínicas empleadas en el estudio de viroma como herramienta para descubrir posibles nuevos virus.

En este apartado se describe el historial clínico de los pacientes empleados en el estudio de la posible presencia de nuevos virus en el viroma respiratorio.

#### 3.3.3.1 Historial clínico paciente 1

Una mujer de 69 años ingresó en el Hospital Universitario Puerta de Hierro (Majadahonda, Madrid) presentando fiebre (superior a 38,4 °C), tos, rinorrea, astenia, dificultad para respirar, dolor abdominal y diarrea. Su historial de viaje reveló un crucero alrededor de los Emiratos Árabes Unidos. Durante la primera evaluación la paciente presentaba taquicardia aunque se encontraba hemodinámicamente estable, la saturación pulsioximétrica estaba al 95% con oxígeno a 3L/minuto y el examen de pecho reveló crujiidos inferiores y medios. Los rayos X de pecho mostraron infiltraciones que afectaban a ambos lóbulos pulmonares inferiores.

El análisis de laboratorio mostró un alto recuento de leucocitos (14.150 cel/ $\mu$ L; 93% neutrófilos), hemoglobina 10.3 g/dl, el tiempo de tromboplastina parcial activada fue de 47.6 segundos, fibrinógeno a 1199 mg/dl, urea a 175 mg/dl; creatinina a 6.1 mg/dl y un aumento aislado de aspartato aminotransferasa.

Se sospechaba una posible infección causada por el CoV-MERS, por ello la paciente fue aislada en una habitación de presión negativa y las precauciones adecuadas fueron implementadas para trabajadores y visitantes.

Su ingreso en UCI se complicó con un síndrome de dificultad respiratoria, lesión renal aguda y deterioro de la función del hígado asociado a coagulopatía por ello se le aplicó un tratamiento

con ceftriaxone, levofloxacin y oseltamivir. Un ecocardiograma transtorácico junto con un ultrasonido abdominal no mostró nada anormal.

Los test de punto de cuidado inicial incluyeron análisis de antígeno de neumococo en exudado nasofaríngeo y de *Legionella* en orina que resultaron negativos.

Durante su estancia en la UCI permaneció sin fiebre requiriendo intubación y ventilación mecánica durante ocho días y diálisis durante dos semanas.

El síndrome respiratorio y el fallo renal y hepático mejoró y la paciente fue dada de alta tras 26 días de tratamiento.

Durante los primeros 8 días de ingreso fueron tomadas 12 muestras y enviadas al Laboratorio de Virus Respiratorios y Gripe del Centro Nacional de Microbiología (Tabla 5).

Identificación	Día de toma	Tipo de muestra
15016	Día 1	SP
15017		ENF
15025		ANF
15027	Día 2	SP
15028		Heces
15039		ATB
15040	Día 3	ENF
15065		ABA
15066	Día 4	ABA
15067		LBA
15078	Día 7	ENF
15093		ATB

**Tabla 5:** Relación de muestras tomadas y día de toma de la muestra de la paciente 1

### 3.3.3.2 Historial clínico “paciente 2”

Varón de 24 años que fue ingresado en un hospital terciario de Madrid presentando síntomas compatibles con una infección respiratoria aguda, como fiebre, tos, rinorrea y dificultad para respirar. Su historial de viaje indicaba una visita a Oriente Próximo por lo que se activó el protocolo de alerta ante la sospecha de infección por CoV-MERS.

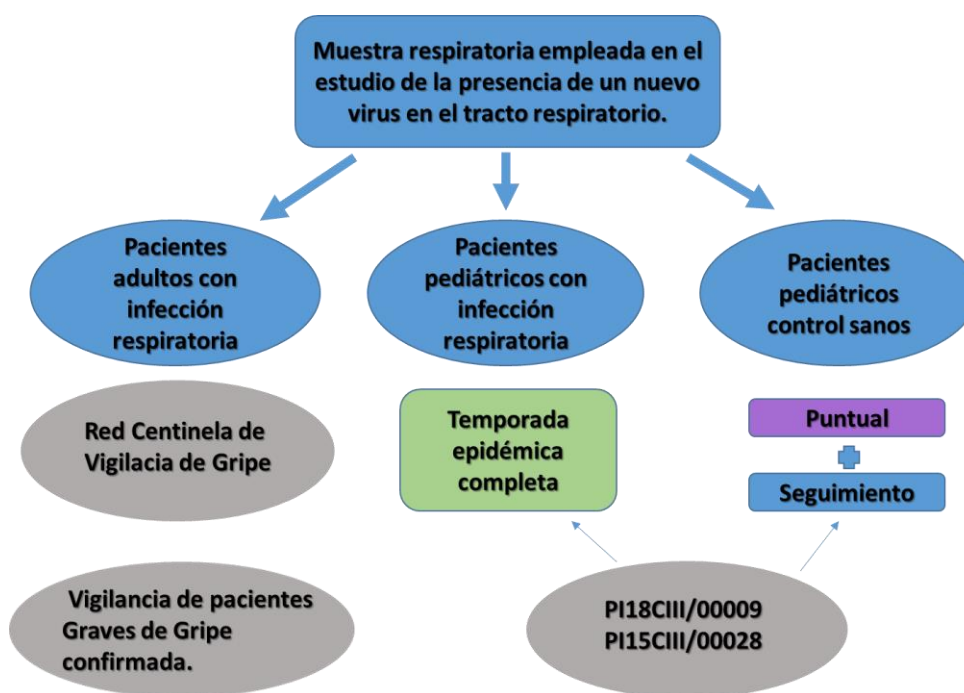
Fueron tomadas 3 muestras, 2 muestras respiratorias y una muestra de orina que fueron enviadas al Laboratorio de Virus Respiratorios y Gripe del Centro Nacional de Microbiología (Tabla 6).

Identificación	Tipo de muestra clínica
162353	ENF
162355	SP
162354	Orina

**Tabla 6:** Identificación y tipo de las muestras del paciente 2

### 3.3.3.3 Características de las muestras con infección respiratoria empleadas en un estudio poblacional sobre la presencia de un posible nuevo virus respiratorio en pacientes adultos y pediátricos

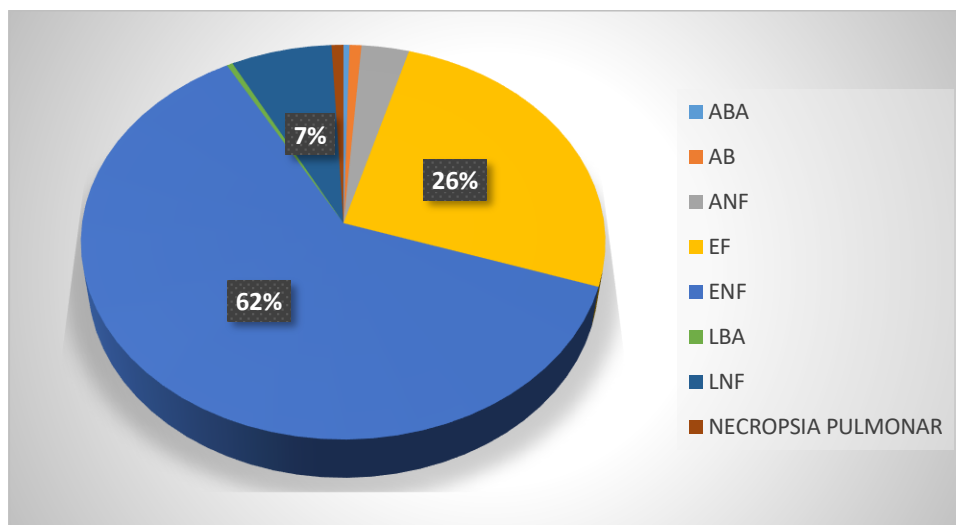
Con el fin de caracterizar la presencia de un posible nuevo virus en el tracto respiratorio se diseñó un estudio en el que fueron empleadas las muestras respiratorias de tres grupos diferentes de pacientes (Figura 3).



**Figura 3:** Esquema de los grupos de pacientes empleados en el estudio de presencia del nuevo virus en el tracto respiratorio y del origen de las muestras tomadas en cada grupo de análisis

El grupo 1 estaba formado por pacientes adultos con infección respiratoria. Estas muestras respiratorias llegaban al Laboratorio de Virus Respiratorios y Gripe del Centro Nacional de Microbiología para la caracterización virológica procedente de pacientes seleccionados por los médicos centinela incluidos en la Red Centinela de Vigilancia de la Gripe, así como de pacientes graves hospitalizados con gripe confirmada en el laboratorio de origen dentro de la Vigilancia de pacientes Graves de Gripe confirmada. En este grupo fueron analizadas 368 muestras respiratorias de pacientes mayores de 16 años. De estas 368 muestras, se emplearon ocho tipos

de muestra respiratoria diferentes, como refleja la Figura 4 el exudado nasofaríngeo es el tipo de muestra más analizado representando el 62% del total de las muestras respiratorias empleadas.

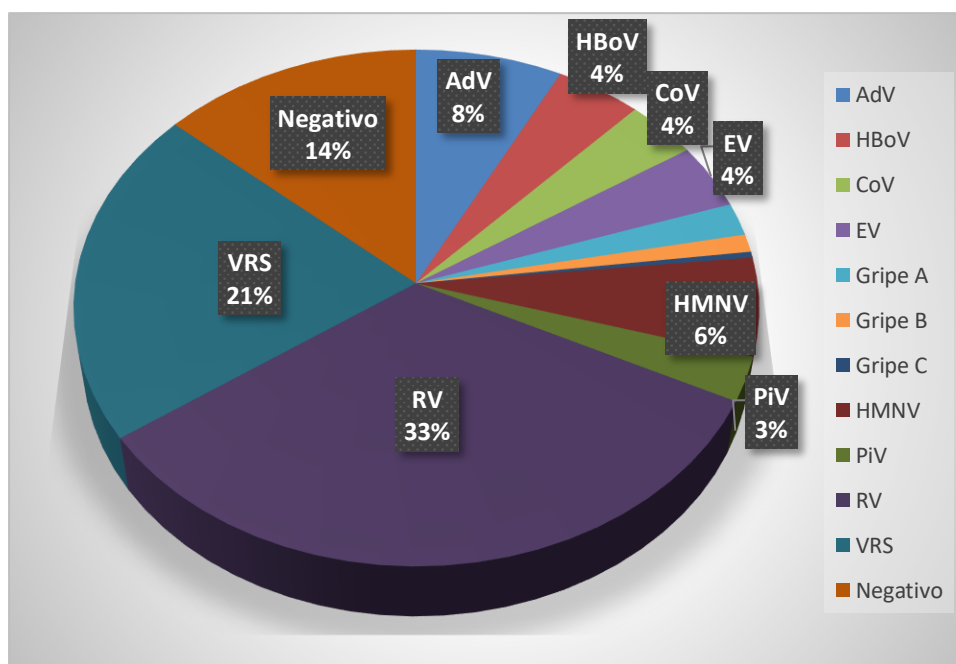


**Figura 4:** Abundancia del tipo de muestra respiratoria empleada en el estudio de la presencia del nuevo virus respiratorio en adultos.

El grupo 2 estaba compuesto por muestras respiratorias pediátricas de pacientes con infección respiratoria. Fueron analizados 361 aspirados nasofaríngeos de pacientes pediátricos con infección respiratoria recibidos durante la temporada epidémica 2017-2018 procedente de los Hospitales Universitarios Severo Ochoa (Leganés) y La Paz (Madrid).

Para el diagnóstico de los pacientes del grupo 2 fueron empleadas las baterías de PCRs de detección de virus respiratorios disponibles en el laboratorio: PCR múltiple en tiempo real Gripe ABC [158], PCR múltiple en tiempo real BRQ [158] y PCR múltiple en tiempo real Respiratorios 2 [159].

Como se describe en la Figura 5 el virus que fue detectado en el mayor número de pacientes fue RV (33%) seguido del VRS (14%). A pesar del predominio de RV y VRS, se detectó la presencia de virus respiratorios diferentes en las 361 muestras analizadas. Durante el análisis de resultados de las PCRs de diagnóstico respiratorio se detectó un 14% de muestras negativas, estas muestras negativas no guardan relación con las muestras control de pacientes sanos.



**Figura 5:** Distribución de la presencia de virus respiratorios en las muestras respiratorias pediátricas empleadas en el grupo 2 del estudio de presencia del nuevo virus en el tracto respiratorio.

El grupo 3 estaba compuesto por las 61 muestras respiratorias del grupo control sano empleada previamente en el estudio de viroma y descritas en el apartado 3.3.2 de Materiales.

### 3.3.4 Características de las muestras empleadas en el desarrollo y aplicación de técnicas NGS en el estudio de los virus gripales y su vigilancia epidemiológica

#### 3.3.4.1 Muestras empleadas en la estandarización de las técnicas NGS para el estudio de los virus gripales.

Para la estandarización y desarrollo de los diferentes métodos empleados en el flujo de trabajo para obtener la secuencia completa del virus de la gripe fueron usados 16 virus previamente caracterizados en el Laboratorio de Virus Respiratorios y Gripe correspondientes a la temporada epidémica 2016-2017 (Tabla 7).

Estos virus fueron detectados en muestras respiratorias procedentes de la Red Centinela de Vigilancia de la Gripe y analizados por la PCR múltiple en tiempo real Gripe ABC [158].

De las 16 cepas empleadas 13 fueron virus de la gripe tipo A subtipo H3N2 y 3 fueron virus de la gripe tipo B linaje Victoria.

Tipo de muestra	Cepa de virus de la gripe	Tipo y subtipo
ENF	<b>A/Navarra/3251/2016</b>	A/H3N2
ANF	<b>A/CastillaLaMancha/3399/2016</b>	A/H3N2
LNF	<b>A/Cantabria/726/2017</b>	A/H3N2
Aislado	<b>A/Extremadura/390/2017</b>	A/H3N2
ENF	<b>A/Ceuta/170026/2016</b>	A/H3N2
ENF	<b>A/Cantabria/1387/2017</b>	A/H3N2
ENF	<b>A/Extremadura/1026/2017</b>	A/H3N2
ENF	<b>A/CastillaLaMancha/769/2017</b>	A/H3N2
LNF	<b>A/Cantabria/730/2017</b>	A/H3N2
Aislado	<b>A/LaRioja/170884/2016</b>	A/H3N2
ENF	<b>A/Galicia/2444/2017</b>	A/H3N2
ANF	<b>A/Galicia/2467/2017</b>	A/H3N2
ENF	<b>A/Baleares/2479/2017</b>	A/H3N2
ENF	<b>B/Galicia/2406/2017</b>	B/Victoria
ANF	<b>B/Galicia/2407/2017</b>	B/Victoria
ENF	<b>B/CastillaLaMancha/2439/2017</b>	B/Victoria

**Tabla 7.** Aislados de gripe, tipo y subtipo de los virus y tipo de muestra de la que proceden los virus empleados en las diferentes estandarizaciones de los protocolos de WGS de gripe

#### 3.3.4.2 Historial clínico del “Paciente 1” y del “Paciente 2” en el estudio de pacientes pediátricos con gripe por WGS

##### Paciente 1

Niña de 7 años que acude a consulta en su Centro de Salud refiriendo fiebre, catarro y otalgia. Inicialmente se prescribió medicación sin antibióticos. Tras cuatro días sin mejora el tratamiento fue modificado y se optó por añadir antibióticos.

Una semana después de la primera visita al centro de salud la paciente acude a urgencias con un cuadro más complejo presentando signos neurológicos e hipotensión. Tras la realización de pruebas diagnósticas se descarta afectación meníngea y regresa a su domicilio. Horas después sufre un síncope en el domicilio y es trasladada a la UCI pediátrica del Hospital de Burgos donde fallece a las pocas horas del ingreso.

El diagnóstico clínico virológico fue realizado en el Centro Nacional de Gripe de Valladolid. El análisis por PCR de las muestras detectó una gripe B perteneciente al linaje Yamagata. Desde el

Centro Nacional de Gripe de Valladolid se enviaron al Laboratorio de Virus Respiratorios y Gripe del Centro Nacional de Microbiología 12 muestras adicionales con el fin de obtener el genoma completo del virus de la gripe presente en la muestra.

<b>Muestras Paciente 1</b>	<b>Identificación</b>
<b>Sangre</b>	181425
<b>Líquido pleural</b>	181426
<b>Bazo</b>	181427
<b>Pulmón</b>	181428
<b>Pulmón</b>	181429
<b>Pulmón</b>	181430
<b>Pulmón</b>	181431
<b>Pulmón</b>	181432
<b>Hisopo faríngeo</b>	181433
<b>Hisopo nasofaríngeo</b>	181434
<b>Oído derecho</b>	181435
<b>Oído izquierdo</b>	181436

**Tabla 8:** Identificación y procedencia de las muestras de la Paciente 1 para análisis WGS de virus de la gripe.

## Paciente 2

Niño de 4 años, hermano de la Paciente 1, que presentó un cuadro respiratorio infeccioso con otalgia y fiebre siendo referido al Hospital Clínico de Valladolid. Debido al desenlace fatal de la “Paciente 1” se le realiza en el Centro Nacional de Gripe de Valladolid una caracterización virológica extensa. El análisis virológico detecta una coinfección gripal por virus de la gripe B del linaje Yamagata y gripe del tipo A subtipo H3N2. Desde el Centro Nacional de Gripe de Valladolid fue enviado un exudado nasofaríngeo y dos cultivos celulares infectados por los virus de la gripe presentes en la muestra del “Paciente 2” para el análisis de la secuencia de genoma completo de estos virus gripales.

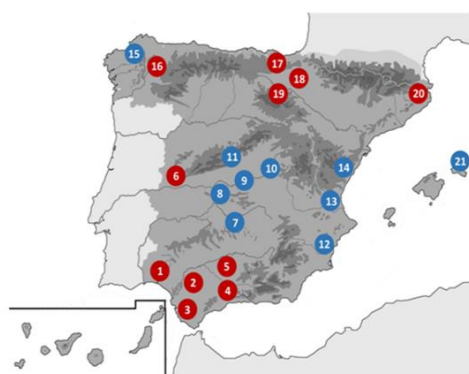
<b>Muestras “Paciente 2”</b>	<b>Identificación</b>
<b>ENF</b>	181434B
<b>Cepa</b>	181434B-SIAT
<b>Cepa</b>	181434B-MDCK

**Tabla 9:** Identificación y tipo de muestra del “Paciente 2” para el análisis WGS de virus de la gripe



### 3.3.5 Muestras de campo empleadas en el estudio de adenovirus de diferentes especies de murciélagos de la Península Ibérica.

Las muestras para el análisis de presencia de adenovirus en murciélagos fueron tomadas entre 2004 y 2016 en varias campañas de campo realizadas en diferentes regiones de la Península Ibérica (Figura 6), durante el desarrollo de los diferentes proyectos de investigación VIROBAT (SAF2006-12784-C02, SAF2009-09172, SAF2013-47149P y SAF2017-89355-P).



**Figura 6:** Mapa de la distribución geográfica de los lugares de captura de los murciélagos analizados. 1. Huelva, 2. Sevilla, 3. Cádiz, 4. Málaga, 5. Córdoba, 6. Cáceres, 7. Ciudad Real, 8. Toledo, 9. Madrid, 10. Guadalajara, 11. Segovia, 12. Alicante, 13. Valencia, 14. Castellón, 15. A Coruña, 16. Lugo, 17. Vizcaya, 18. Navarra, 19. La Rioja, 20. Girona, 21. Menorca. Los círculos rojos representan los lugares con muestras positivas para mastadenovirus de murciélagos.

Los métodos para la obtención de muestras siguieron las regulaciones y procedimientos éticos de la Sociedad Española para la Conservación y el Estudio de los Murciélagos (SECEMU). Para el estudio virológico se tomaron dos tipos de muestras: exudado orofaríngeo (EOR) y heces (H). Las muestras fueron tomadas de 27 especies diferentes de murciélagos pertenecientes a 3 familias. Se analizaron un total de 1717 muestras (Tabla 10)

### 3.3.6 Muestras de murciélagos empleadas en la puesta a punto de las técnicas de secuenciación de nueva generación para adenovirus

Parte de las muestras de campo analizadas fueron empleadas para la estandarización de técnicas para la secuenciación de genoma completo de AdV

Las muestras empleadas fueron 8 EOR de los cuales 3 pertenecían a la especie *Pipistrellus kuhlii* y los 5 restantes pertenecían a la especie *Pipistrellus pygmaeus* (Tabla 11). Estas 8 muestras resultaron positivas a AdV tras la realización de las PCRs de hexón y polimerasa de AdVs.

Familia	Especies	EOR	H	Años de captura
<b>Vespertilionidae</b>	<i>Barbastella barbastellus</i> (Bba)	38	4	2007,2008
	<i>Eptesicus isabellinus</i> (Eis)	0	8	2004,2007
	<i>Eptesicus serotinus</i> (Ese)	0	14	2003,2007
	<i>Hypsugo savii</i> (Hsa)	31	26	2007
	<i>Myotis alcathoe</i> (Mal)	0	1	2007
	<i>Myotis bechsteinii</i> (Mbe)	18	2	2007
	<i>Myotis blythii</i> (Mbl)	29	0	2004
	<i>Myotis capaccinii</i> (Mca)	15	0	2004
	<i>Myotis daubentonii</i> (Mda)	63	41	2004,2007
	<i>Myotis emarginatus</i> (Mem)	56	0	2004,2007,2008
	<i>Myotis escalerae</i> (Mes)	13	0	2004,2007
	<i>Myotis myotis</i> (Mmy)	79	1	2004,2007
	<i>Myotis mystacinus</i> (Mmt)	2	8	2007
	<i>Myotis nattereri</i> (Mna)	36	3	2007
	<i>Nyctalus noctula</i> (Nno)	122	0	2007
	<i>Nyctalus lasiopterus</i> (Nylas)	139	40	2007
	<i>Nyctalus leisleri</i> (Nyle)	19	26	2007
	<i>Pipistrellus kuhlii</i> (Pku)	350	4	2007
	<i>Pipistrellus pipistrellus</i> (Ppi)	29	4	2007
	<i>Pipistrellus pygmaeus</i> (Ppy)	36	120	2007
	<i>Plecotus auritus</i> (Pau)	11	8	2007
	<i>Plecotus austriacus</i> (Pas)	10	6	2004,2007
<b>Miniopteridae</b>	<i>Miniopterus schreibersii</i> (Msc)	152	2	2002,2004, 2007
<b>Rhinolophidae</b>	<i>Rhinolophus euryale</i> (Rheu)	49	0	2004,2007,2008
	<i>Rhinolophus ferrumequinum</i> (Rfe)	90	3	2004,2007,2008
	<i>Rhinolophus hipposideros</i> (Rhi)	4	4	2007,2008
	<i>Rhinolophus mehelyi</i> (Rme)	1	0	2007
<b>Total</b>		<b>1392</b>	<b>325</b>	

**Tabla 10:** Familia, especie, tipo de muestra y año de captura de los murciélagos empleados en el estudio de adenovirus de murciélagos de la Península Ibérica

Identificación	Muestra	Especie
<b>Pku2 160622</b>	EOR	<i>Pipistrellus kuhlii</i>
<b>Pku15 160518</b>	EOR	<i>Pipistrellus kuhlii</i>
<b>Pku40 160517</b>	EOR	<i>Pipistrellus kuhlii</i>
<b>Ppy14 160725</b>	EOR	<i>Pipistrellus pygmaeus</i>
<b>Ppy15 160622</b>	EOR	<i>Pipistrellus pygmaeus</i>
<b>Ppy23 160628</b>	EOR	<i>Pipistrellus pygmaeus</i>
<b>Ppy29 160622</b>	EOR	<i>Pipistrellus pygmaeus</i>
<b>Ppy15 181106</b>	EOR	<i>Pipistrellus pygmaeus</i>

**Tabla 11:** Identificación, tipo de muestra y especie de los murciélagos empleados en el estudio para WGS de adenovirus

## 3.4. Oligonucleótidos

### 3.4.1 Oligonucleótidos publicados previamente a esta tesis

En la Tabla 12 se indican los nombres de los oligonucleótidos cebadores, sus secuencias, la técnica de PCR en la que son empleados y la referencia bibliográfica del artículo que describe dichos oligonucleótidos empleados.

PCR	Nombre	Secuencias	REF
<b>Hexón (Adenovirus)</b>	gADV1s	5' GTIGAYTGCAIGACAGRAAYACIGA 3'	[2]
	gADV1r	5' TTTIAGICKIGTRAAISWCCAICC 3'	
	HADV2+	5' AGITAYTTYWGIATGTGGAA 3'	
	panADV	5' TGRTCRTTGGTITCRTTICKIAGCAT 3'	
<b>PCR multisegmento universal de gripe</b>	unifluF	5'-ACGACGGGCGACAIAGCARAAGC-3'	[18]
	unifluR	5'-ACGACGGGCGACAAGTAGWAACA-3'	
	uniflu	5'-IAGCARAAGC -3'	
<b>PCR multisegmento Gripe A</b>	Uni12/Inf-1	5'-GGG GGG AGC AAA AGC AGG-3'	[16]
	Uni12/Inf-3	5'-GGG GGG AGC GAA AGC AGG-3'	
	Uni13/Inf-1	5'-CGG GTT ATT AGT AGA AAC AAG G-3'	
<b>PCR multisegmento Gripe B</b>	B-PBs-UniF	5'-GGGGGGGAGCAGAAGCGGAGC-3'	[17]
	B-PBs-UniR	5'-CCGGGTATTAGTAGAAACACGAGC-3'	
	B-PA-UniF	5'-GGGGGGGAGCAGAAGCGGTGC-3'	
	B-PA-UniR	5'-CCGGGTATTAGTAGAAACACGTGC-3'	
	B-HANA-UniF	5'-GGGGGGGAGCAGAAGCAGAGC-3'	
	B-HANA-UniR	5'-CCGGGTATTAGTAGTAACAAGAGC-3'	
	B-NP-UniF	5'-GGGGGGGAGCAGAAGCACAGC-3'	
	B-NP-UniR	5'-CCGGGTATTAGTAGAAACAACAGC-3'	
	B-M-Uni3F	5'-GGGGGGGAGCAGAAGCACGCACTT-3'	
	B-Mg-Uni3F	5'-GGGGGGGAGCAGAAGCAGGCACTT-3'	
	B-M-Uni3R	5'-CCGGGTATTAGTAGAAACAACGCACTT-3'	
	B-NS-Uni3F	5'-GGGGGGGAGCAGAAGCAGAGGATT-3'	
	B-NS-Uni3R	5'-CCGGGTATTAGTAGTAACAAGAGGATT-3'	
<b>Amplificación independiente de secuencia</b>	FR26RV-N	5' GCC GGA GCT CTG CAG ATA TCNNNNNN 3'	[12]
	FR20RV	5' GCC GGA GCT CTG CAG ATA TC 3'	
<b>PCR múltiple tiempo real BRQ</b>	HMPVrtS	5'-AGAGCTGAAAGAITTTGTGAGCAA-3'	[158]
	HMPVrtA	5'-ATTARGTCCAAGWATATTGCTGGTG-3'	

	RSAB1F	5'-TTCTTGTTAGGTGTWGGATCTGCAAT-3'	
	RSAB1A	5'-GTTACACCTGCATTRACACTRAATTC-3'	
	BoVB-rtS-13	5'-GAC GAR GAA ATG CTT TCT GCT G-3'	
	BoVB-rtA-13	5'-CAC CGT TAT CAA GWG GAT TAA ATG G-3'	
	ADVrtS-11	5'-CAG TGG YCR TAC ATG CAC ATC-3'	
	ADVrtA-11	5'-ACC GTG GGR TTY CTA AAC TTR TT-3'	
<b>PCR múltiple tiempo real respiratorios 2</b>	HER1	5'-CTCCGGCCCCTGAATRYGGCTAA-3'	[159]
	HER4	5'-CTGTGTTGAWACYTGAGCICCCA-3'	
	PIV-S	5'-TCGMAAAATAYTGTCTTMAITGGAGAT-3'	
	PIV-A	5'-GGCTTGATTRTCTCCYGIACC-3'	
<b>PCR múltiple Gripe ABC</b>	InfA-F	5'-GAC CRA TCC TGT CAC CTC TGA C-3'	[158]
	InfA-R	5'-AGG GCA TTY TGG ACA AAK CGT CTA-3'	
	NPB+10	5'-AAC AAA AGG CTC TAG TTG ATC AAG TGA T-3'	
	NPB-11	5'-CTT CAT AGG CAG CTC CGA AGC-3'	
	NPCS+10	5'-CAA ACT ACA ATG ATG CCA GAG AAG C-3'	
	NPC-11	5'-GAA GTC TTC ATA AGC AAG TCC AAA GC-3'	
<b>PCR MERS-CoV</b>	ORF1a_F	5'-CCACTACTCCCATTTCGTCAG-3'	[160]
	ORF1a_R	5'-CAGTATGTGTAGTGCGCATATAAGCA-3'	
	upE-F	5'-GCAACGCGCGATTTCAGTT-3'	
	upE-R	5'-GCCTCTACACGGGACCCATA-3'	

**Tabla 12:** Secuencia, nombre y PCR en la que se emplean estos oligonucleótidos cebadores procedentes de trabajos previamente publicados

### 3.4.2 Oligonucleótidos diseñados en esta tesis

En la Tabla 13 se indican los nombres de los oligonucleótidos cebadores, sus secuencias, la técnica de PCR en la que se emplean y el tamaño del producto de amplificación generado de las diferentes PCRs diseñadas y estandarizadas en este trabajo con los objetivos de 1. amplificar un fragmento del gen que codifica la ADN polimerasa ADN dependiente de Adenovirus y 2. estudiar y caracterizar un nuevo virus circular (HRCiV).

PCR	Nombre	Secuencias	Tamaño
<b>Polimerasa (Adenovirus)</b>	PolF	5'-GTIGCRAAIGAICCRTAGAGGGC-3'	450 pb
	PolR	5'-GTTTAYGAYATITGYGGMATGTAYGC-3'	
	PolF2	5'-AAIGAICCRTAGAGGGCRRTTKGA-3'	
<b>REP de HRCiV</b>	HRCiV2F	5'-AGAATACTGCCAAAAAGACGGG-3'	381 pb
	HRCiV2R	5'-AGCTTTTCTGAAGTCATCAATTATTGC-3'	
<b>Genoma completo HRCiV</b>	HRCiVT1	5'-GCTATGTTAACCGACTGGAGC-3'	1500 pb
	HRCiVT4	5'-ATAATCCCATACTTACGCCGG-3'	1700 pb
	HRCiVT5	5'-GATGAACTGCTCTATTCCTTGG-3'	
	HRCiVT8	5'-CCAGTTTAAGCTTATCTTCATCC-3'	

**Tabla 13.** Nombre, secuencia del oligonucleótido cebador empleado y tamaño del producto de amplificación de las PCRs diseñadas en esta tesis.

### 3.4.3 Sondas TaqMan empleadas en las diferentes técnicas de PCR tiempo real

En la Tabla 14 se describen las secuencias específicas de las sondas utilizadas en las diferentes técnicas de PCR múltiple en tiempo real empleadas en esta tesis.

PCR	Nombre de sonda	Secuencias
<b>PCR múltiple tiempo real BRQ</b>	MPV-Cy5	5'Cy5-TG TTG TGC G/TAO/G CAR TTT TCA GAC AAY GC-3' IABRQSp
	vicADVrt-11	5'VIC-TGG TGC ART TYG CC-3'MGB
	RSVAB-HEX	5'HEX-TC TAA CAC TT TGC TGG TTA ARA CAC TGA CC-3'
	BoV-TYE7	5'TYE705-GAG GGT GGG TGC TKC CTG GAT AC-3'
<b>PCR múltiple tiempo real respiratorios 2</b>	PIV1	5'6-FAM-CT GTT AAA GT TGG TGT CAG AGT GTC AGC -3'IABkFQ
	PIV2	5'6-FAM-CA TTT TCT GA CAT AGG CCT TCA ATA CCT CC-3'IABkFQ
	PIV3	5'6-FAM-CT CGT CTT GA AGG AAG TAC AAT CTA TGT AGG -3'IABkFQ
	PIV4	5'6-FAM-CC AAG AGG AG GTA TWG AAG GTC TGT GCC-3'IABkFQ
	EV-RIN	5'HEXAAACACGGACACCCAAAGTAGTYGGTYC-3'
<b>PCR múltiple Gripe ABC</b>	FluA Fam	5'6-FAM-TGC AGT CCT CGC TCA CTG GGC ACG-3'BHQ1
	FluB Vic	5'VIC-TGG CAA CAC CTG TTT C-3'MGB
	FluC Cy5	5'Cy5-GGC TGC CAG ACT TAC TCTG AAA GAT GAG-3'
<b>PCR MERS-CoV</b>	ORF1a	3'FAM-TTGCAAATTGGCTTGCCCCACT-3'TAMRA
	UpE	5'FAM-CTCTTCACATAATCGCCCCGAGCTCG-3'TAMRA

**Tabla 14:** Sondas utilizadas en las técnicas de PCR múltiple tiempo real empleadas en esta tesis

## 3.5. Reactivos

### 3.5.1. Reactivos comerciales

- ADNasa I (Roche diagnostics GmbH, Mannheim, Alemania)
- Agarosa *SEAKem LE* (Cambrex, Maryland, EEUU)
- *AmpliTaq* (Applied Biosystems, Branchburg, NJ, EEUU).
- ARNasa A (Roche diagnostics GmbH, Mannheim, Alemania)
- Azul de Bromofenol (Bio-Rad Laboratories, California, EEUU)
- Bolas *Ampure XP* (Beckman Coulter, Brea, California, EEUU)
- Chip de alta sensibilidad *Agilent* (Agilent technologies, Waldbronn, Alemania)
- Citrato Sódico (Sigma, Sigma-Aldrich Chemie GmbH, Munich, Alemania)
- Dimetilsulfóxido (Sigma, Sigma-Aldrich Chemie GmbH, Munich, Alemania)
- dNTPs (GE Healthcare, Buckinghamshire, Reino Unido)
- Enzima de restricción *NcoI* (New England Biolabs Ltd, Reino Unido)
- Etanol (Applichem, Darmstadt, Alemania)
- Filtro de centrífuga *Millipore* de poro 0,45 µm (Merck Millipore, Cork, Irlanda)
- Filtro de centrífuga *Millipore* de poro 5 µm (Merck Millipore, Cork, Irlanda)
- *Gel Red* (Biotium Corporate, California, EEUU)
- Glicógeno (Sigma, Sigma-Aldrich Chemie GmbH, Munich, Alemania)
- H<sub>2</sub>O *PCR grade* (Roche diagnostics GmbH, Mannheim, Alemania)
- Isopropanol (Merck Millipore, Cork, Irlanda)
- *Klenow exopolimerasa 3'->5'* (New England Biolabs Ltd, Reino Unido)
- *Lightcycler 480 probes master*(Roche Diagnostics, Mannheim, Alemania)
- Marcador de peso molecular de 100 pb: (Promega corporation, Madison, EEUU)
- Marcador de peso molecular de 1000 pb: (GeneON GmbH, Ludwigshafen am Rhein, Alemania)
- *MyTaq Red Mix* (Bioline Reagents, Reino Unido)
- Nucleasa S7 (Roche Diagnostics, Mannheim, Alemania)
- Proteasa K (Qiagen GmbH, Hilden, Alemania)
- *Sarkosyl* (Sigma, Sigma-Aldrich Chemie GmbH, Munich, Alemania)
- Tampón de nucleasas 10x (Roche Diagnostics, Mannheim, Alemania)

- Tampón de restricción *NEB* (New England Biolabs Ltd, Reino Unido)
- Tampón *Klenow* exopolimerasa 3'-5' (New England Biolabs Ltd, Reino Unido)
- Tampón *Taq* polimerasa con  $(\text{NH}_4)_2\text{SO}_4$  y  $\text{MgCl}_2$  20mM (Thermo Fisher, Reino Unido)
- Tiocianato de guanidinio (Sigma, Sigma-Aldrich Chemie GmbH, Munich, Alemania)

### 3.5.2 Kits comerciales

- *Big Dye Terminator v.3.1* (Applied Biosystem, Foster City, EEUU)
- *Illustra TempliPhi Amplification Kit* (GE Healthcare, Illinois, EEUU)
- Indices *Nextera XT V2* (Illumina, San Diego, EEUU)
- Indices *SeqCap Adapter A* (Roche diagnostics, Indianapolis, EEUU)
- Indices *SeqCap Adapter B* (Roche diagnostics, Indianapolis, EEUU)
- Panel de captura *HyperCap Target Enrichment* (Roche diagnostics, Indianapolis, EEUU)
- Preparación de librerías *Kapa Hyper Plus* (Kapa Biosystems, Cape Town, Sudáfrica)
- Preparación de librerías *Nextera XT* (Illumina, San Diego, EEUU)
- *QIAamp Mini Elute Virus* (Qiagen, Valencia, CA, EEUU)
- *QIAquick PCR purification* (Qiagen, Valencia, CA, EEUU)
- *Quantifluor* (Promega corporation, Madison, EEUU)
- Reactivos *MiSeq v2 300 ciclos* (Illumina, San Diego, EEUU)
- *SuperScript III First Strand Synthesis System* (Invitrogen, California, EEUU)
- *Superscript III One-step RT-PCR System con Platinum Taq High Fidelity* (Invitrogen, California, EEUU)
- *SuperScript III PlatinumOne-Step RT-PCR System* (Invitrogen, California, EEUU)

### 3.5.3 Medios y reactivos de cultivo celular

- DMEM, *Dulbecco's modified Eagle's medium*, (Sigma-Aldrich, Steinheim, Reino Unido)
- L-Glutamina (Lonza, Verviers, Bélgica)
- Penicilina- estreptomycin (Lonza, Verviers, Bélgica)
- Suero fetal bovino (Gibco by Life Technologies, Grand Island, EEUU)
- Tripsina (Gibco by Life Technologies, Grand Island, EEUU)

### 3.5.4 Equipo y aparatos

- Bioanalizador *Agilent* (Agilent technologies, Waldbronn, Alemania)
- Estufa de 37°C con CO<sub>2</sub>
- Extractor automático *QIAcube* (Qiagen, Valencia, CA, EEUU)
- Fluorímetro *Quantus* (Promega corporation, Madison, EEUU)
- Minifuga *Biofuge pico 21* (Heraeus, Heraeus Holding GMBH, Hanau, Alemania)
- Secuenciador *ABI PRISM 3700* (Applied Biosystem, Foster City, EEUU)
- Secuenciador *Illumina MiSeq* (Illumina, San Diego, EEUU)
- Secuenciador *MinION* (Oxford Nanopore Technologies, Oxford, Reino Unido)
- Sistema de electroforesis horizontales *BioRad* (Bio-Rad Laboratories, California, EEUU)
- Termociclador C1000 (Bio-Rad Laboratories, California, EEUU)
- Termociclador CFX6 (Bio-Rad Laboratories, California, EEUU)

### 3.5.5 Programas informáticos

- *A5-miseq* (<http://sourceforge.net/projects/ngopt>)
- *ABRicate* (<https://github.com/tseemann/abricate>)
- *BLAST*: Herramienta de búsqueda de alineamientos locales básicos (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>)
- *Bowtie2* v.2.1.6 (<http://bowtie.cbcb.umd.edu>.)
- *Bwasw* v.0.7.9a-r786 ([bio-bwa.sourceforge.net](http://bio-bwa.sourceforge.net))
- *Chromas PRO* v.1.5 (Technelysium Ltd)
- *CLUSTAL X* v.2.0 ([www.clustal.org](http://www.clustal.org)).
- *Cutadapt* v.1.7 (<http://code.google.com/p/cutadapt/>)
- *DIAMOND* (<http://www.diamondsearch.org/>)
- Editor de secuencias *EditSeq* (Lasergene-DNAStar Software)
- *FastQC* (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc>)
- *Galaxy* (<https://usegalaxy.org/>)
- *IDBA-UD* ([www.cs.hku.hk/~alse/idba\\_ud](http://www.cs.hku.hk/~alse/idba_ud))
- *INSaFLU* (<https://insaflu.insa.pt/>)
- *IRMA* (<https://wonder.cdc.gov/amd/flu/irma/irma.html>)
- *LABEL* ([label.phiresearchlab.org](http://label.phiresearchlab.org))



- *Macaw* (National Center for Biotechnology Information, National Library of Medicine, Maryland, EEUU)
- *MEGA 5.2* ([www.megasoftware.net](http://www.megasoftware.net))
- *Megablast v.2.2.26* (<https://blast.ncbi.nlm.nih.gov>)
- *Modeltest* ([http://darwin.uvigo.es/software/modeltest\\_server](http://darwin.uvigo.es/software/modeltest_server))
- *Oligocalc* (<http://biotools.nubic.northwestern.edu/OligoCalc.html>)
- *ORFfinder* (<https://www.ncbi.nlm.nih.gov/orffinder/>)
- *Picard* (<http://broadinstitute.github.io/picard>)
- *Platanus v.1.2.1* (<http://platanus.bio.titech.ac.jp>)
- *PriceTI* ([derisilab.ucsf.edu](http://derisilab.ucsf.edu))
- *Prinseq-lite v.0.20.4* (<http://prinseq.sourceforge.net/>.)
- *RAP Search2 v. 2.16* (<http://omics.informatics.indiana.edu/mg/RAPSearch2/>)
- *Ray v.2.2* (<http://denovoassembler.sf.net>.)
- *Skewer* (<https://sourceforge.net/projects/skewer>)
- *Snippy* (<https://github.com/tseemann/snippy>)
- *SPAdes* (<http://cab.spbu.ru/software/spades/>)
- *Trimmomatic* (<http://www.usadellab.org/cms/index.php?page=trimmomatic>)
- *Virus TAP* (<https://gph.niid.go.jp/cgi-bin/virustap/index.cgi>)
- Visualizador de geles de agarosa *GelDoc* (Bio-Rad Laboratories, California, EEUU)

### 3.5.6 Bases de datos

- Base de datos GISAID: “Global Initiative on Sharing All Influenza” (<https://www.gisaid.org/>)
- Base de datos GenBank (<https://www.ncbi.nlm.nih.gov/genbank/>)

# MÉTODOS:



## MÉTODOS

### 4.1. Mantenimiento de líneas celulares

Las líneas celulares VERO E-6, L132, Hep2 y MDCK-SIAT (Materiales. Apartado 3.1.) empleadas en esta tesis fueron crecidas en monocapas de células en frascos de 75 cm<sup>2</sup> empleando el medio de crecimiento DMEM suplementado con L-Glutamina, penicilina-estreptomicina y de suero fetal bovino (SFB).

El SFB previamente inactivado a 56°C durante 30 minutos aporta nutrientes y otros componentes que permiten el crecimiento de las células, el aminoácido esencial L-Glutamina aporta el nitrógeno que sirve de fuente de energía al metabolismo celular y el uso de los antibióticos penicilina-estreptomicina a una concentración de 10000 U tiene como fin evitar contaminación por bacterias.

El medio de crecimiento empleado para el mantenimiento de las células VERO E-6, L132 y Hep2 (Materiales, Apartado 3.1) fue un medio DMEM suplementado con 1% de penicilina-estreptomicina, 2% de glutamina y 10% de SFB (Materiales. Apartado 3.5.3). En el caso de las células MDCK-SIAT se empleó un medio DMEM suplementado con 1% de penicilina/estreptomicina, 1% de glutamina y 5% de SFB (Materiales. Apartado 3.5.3). Todas las células se incubaron a 37°C con una saturación de 5% CO<sub>2</sub> para mantener el pH del cultivo.

La tripsina (Materiales. Apartado 3.5.3) fue empleada para el desprendimiento del tapiz en los pases de mantenimiento.

### 4.2. Aislamiento de virus en las líneas celulares VERO-E6, L132, Hep2 y MDCK-SIAT

El aislamiento viral en las líneas celulares empleadas se realizó al 70% de confluencia de las monocapas de células.

Inicialmente las células a emplear fueron lavadas con medio DMEM (Materiales. Apartado 3.5.3) para limpiar los restos de medio de crecimiento de las líneas celulares. Tras los lavados fueron inoculados 200 µl de la muestra clínica de interés, el tiempo de adsorción fue de una hora a 37°C. Tras la retirada del inóculo se efectuaron lavados post-infección con medio DMEM, conteniendo 1% de penicilina-estreptomicina para evitar la contaminación por presencia de bacterias. Finalmente fueron añadidos 2 ml de medio post-infección compuesto por DMEM suplementado

con 1% de antibiótico y glutamina, y 2% de SFB (Materiales. Apartado 3.5.3). Tras la adición del medio post-infección las monocapas celulares fueron incubadas a 37°C con una saturación de 5% CO<sub>2</sub>.

En el caso de las líneas celulares L132 y VERO E-6 se realizaron algunos cambios en el protocolo de trabajo descrito anteriormente con el fin de optimizar el método para obtener mejor rendimiento.

En la línea celular L132 fue realizada una prueba de ajuste empleando medio post-infección con diferente porcentaje de SFB como suplemento o con presencia de tripsina para ver el efecto en la infección del virus circular HRCiV. Los medios empleados fueron medios suplementados con: 1. con SFB; 2. sin SFB; 3. con SFB y 0,2% de tripsina; 4. sin SFB y 0,2% de tripsina; 5. con SFB y 1% de tripsina y 6. sin SFB y 1% de tripsina (Materiales. Apartado 3.5.3).

Para la inoculación de la línea celular VERO E-6 con AdV de murciélagos fueron inoculadas las monocapas celulares con 100 µl y 200 µl de muestra de campo tomada de murciélagos que previamente habían resultado ser positivos a AdV mediante métodos moleculares.

La adsorción se realizó durante una hora a dos temperaturas diferentes: 32°C, temperatura corporal de los murciélagos, y a 37°C siguiendo el protocolo de trabajo habitual de aislamiento viral en esta línea celular.

### **4.3 Tratamiento de muestras clínicas**

Para el procesamiento de las muestras clínicas se siguió el protocolo acreditado de calidad según Norma ISO 15189 (VRP PT 002 02) desarrollado e implementado en el Laboratorio de Virus Respiratorios y Gripe. Las muestras fueron procesadas dentro de las 24 horas posteriores a la toma y a su llegada al laboratorio fueron preparadas 3 alícuotas de cada muestra posteriormente almacenadas a -80°C. La recepción y el procesamiento de muestras se realizaron en áreas de trabajo separadas de las áreas de preparación de reactivos, extracción y amplificación.

Las muestras de heces fueron homogeneizadas en medio DMEM al 10% de suspensión (Materiales. Apartado 3.5.3) y centrifugadas a 6000xg durante 10 minutos. El sobrenadante fue alicuotado y almacenado a -80°C.

Las muestras de campo obtenidas directamente de los animales capturados se introdujeron en tubos preparados de 1 ml de tampón lisis para su posterior almacenamiento a -80°C hasta su análisis.

## 4. 4 Extracción de ácidos nucleicos totales

### 4.4.1 Extracción manual por precipitación alcohólica de ácidos nucleicos totales

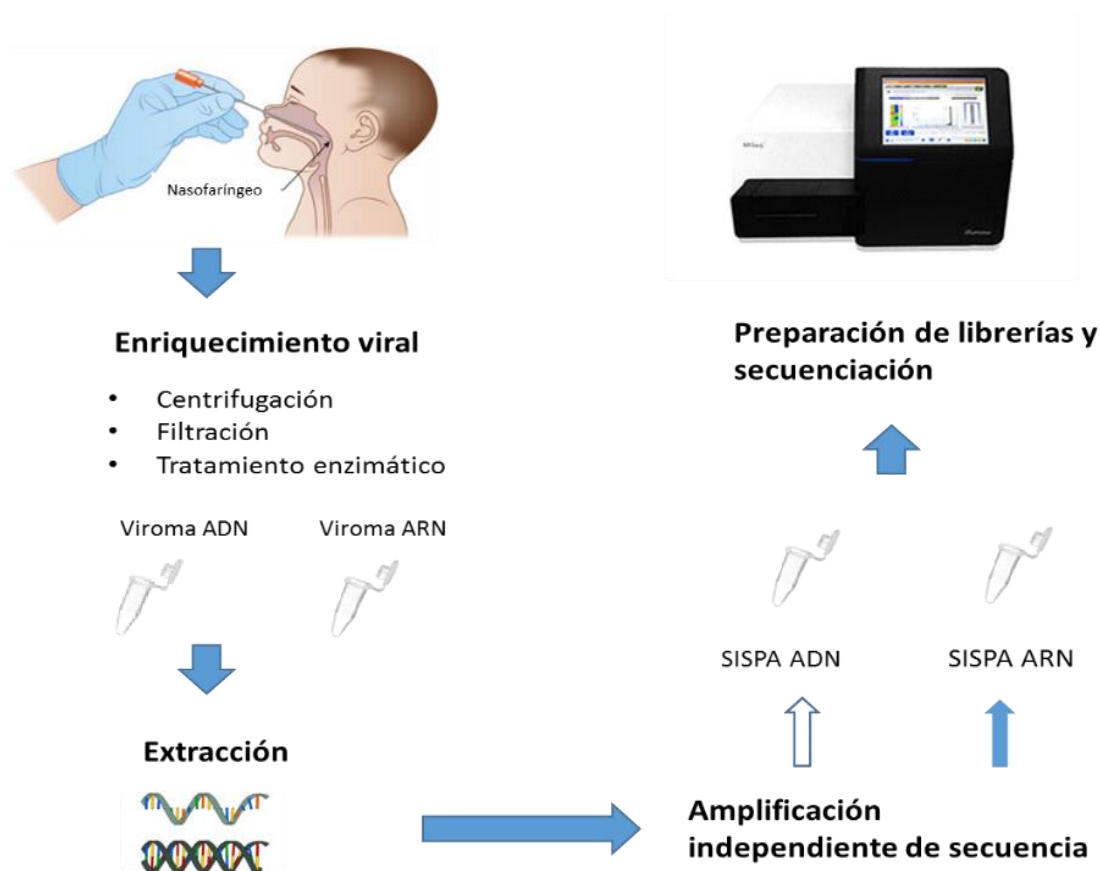
Se emplea una extracción manual con precipitación alcohólica para extraer los ácidos nucleicos totales de las muestras clínicas evitando el uso de columnas de sílice. El protocolo empleado fue descrito en trabajos previos [161] y consistió en la adición de 200 µl de la muestra a extraer a 600 µl de un tampón de lisis compuesto de tiocianato de guanidinio 4M, *sarkosyl* 0,5%, ditioneitol 1mM, citrato sódico 25 mM y glicógeno a 0,1 µg/µl (Materiales. Apartado 3.5.1). Tras la lisis de la muestra se adicionaron 600 µl de isopropanol a -20°C (Materiales. Apartado 3.5.1) para la precipitación de los ácidos nucleicos y posteriormente se realizó una centrifugación durante 10 minutos a 14000xg y 4°C. Una vez retirado el isopropanol se realizó un lavado del precipitado obtenido con etanol al 70% (Materiales. Apartado 3.5.1. Páginas 77-78). Tras una segunda centrifugación de 10 min a 14000xg y 4°C, el etanol fue eliminado y el precipitado seco fue resuspendido en 30 µl de agua libre de ADNasas.

### 4.4.2 Extracción automática de ácidos nucleicos totales

Para la extracción de ácidos nucleicos totales de muestra clínica se siguió el protocolo acreditado de calidad según Norma ISO 15189 (VRP PT 004 03) desarrollado e implementado en el Laboratorio de Virus Respiratorios y Gripe.

La extracción de los ácidos nucleicos se realizó con el extractor automático *QIAcube* (Materiales. Apartado 3.5.4. Página 79). Las muestras que se procesaron en el extractor automático *QIAcube* fueron previamente inactivadas mediante el tampón de lisis comercial procedente del kit *QIAamp Mini Elute Virus* (Materiales. Apartado 3.5.2. Página 79). Las muestras inactivadas fueron transferidas a continuación a unas columnas de extracción que fueron sometidas a centrifugación. Los ácidos nucleicos quedaron retenidos en una membrana de sílice o resina de purificación presente en la parte inferior de la columna de extracción donde fueron realizados varios procesos de lavado para eliminar residuos. Las columnas fueron transferidas finalmente a tubos de elución donde tras la adición de tampón de elución comercial procedente del kit *QIAamp Mini Elute Virus* (Materiales. Apartado 3.5.2. Página 79) se obtuvieron los ácidos nucleicos totales purificados.

## 4.5. Flujo de trabajo para la realización del estudio de viroma



**Figura 7:** Flujo de trabajo para el estudio de viroma.

La preparación de muestras para el estudio de viroma se realiza mediante un conjunto de protocolos relacionados (Figura 7) que incluye desde el enriquecimiento de la proporción de genomas virales en la muestra hasta la preparación de librerías para su secuenciación.

### 4.5.1 Enriquecimiento viral por tres métodos

Uno de los procedimientos de trabajo desarrollados durante esta tesis fue el enriquecimiento mediante diferentes métodos, que tienen como objetivo el aumentar la proporción viral presente en las muestras clínicas de interés. Los métodos empleados fueron: 1. una centrifugación inicial; 2. un filtrado doble y 3. el empleo de enzimas de digestión para el estudio de viroma ADN y ARN.

Las muestras respiratorias que presentaban un mayor grado de viscosidad requirieron una homogenización previa mediante vórtex y disgregado del esputo con jeringuilla. Una vez

fluidificada, la muestra fue centrifugada durante 15 minutos a 6000xg con el fin de depositar fragmentos celulares, y otros materiales de gran tamaño al fondo del tubo.

El sobrenadante resultante de la centrifugación fue sometido a dos filtrados consecutivos usando filtros de centrífuga de 5 µm y otro de 0,45 µm (Materiales. Apartado 3.5.1. Páginas 77-78). Estos filtrados tienen como fin eliminar bacterias y otros elementos de mayor tamaño que los virus.

Posteriormente el material fue dividido en dos alícuotas de 200 µl cada una con el fin de realizar los diferentes tratamientos para el estudio de viroma ADN y ARN. El tratamiento enzimático es específico según el tipo de viroma que queremos enriquecer en la muestra por ello se partió de dos alícuotas diferenciadas para los dos tipos de tratamiento enzimático.

Para la digestión enzimática del material filtrado empleado en el estudio del viroma ADN se preparó una mezcla compuesta de tampón de nucleasas 10x, ADNasa I (500 U/ml), ARNasa A (10 µg/ml) y nucleasa S7 (100 U/ml) (Materiales. Apartado 3.5.1) que fue incubada junto con la muestra durante 45 minutos a 37°C.

Tras esta incubación inicial fue añadida la enzima proteasa K (Materiales. Apartado 3.5.1) e incubada nuevamente a 56°C durante 15 minutos.

El protocolo para la digestión enzimática del material empleado para el estudio de viroma ARN estuvo conformado por una incubación durante 20 minutos a 37°C del tampón de nucleasas 10x, la ADNasa I (500 U/ml) (Materiales. Apartado 3.5.1) y la muestra. Una vez finalizada la digestión se produjo la inactivación de la enzima a 67°C durante 10 minutos.

Tras los tratamientos enzimáticos los ácidos nucleicos de las muestras fueron extraídos de forma automática como se describe en (Métodos. Apartado 4.4.2) empleando el kit *QIAamp Mini Elute Virus Spin* (Materiales. Apartado 3.5.2) en el extractor automático *Qiacube* (Materiales. Apartado 3.5.4). En el caso de la extracción de las muestras enriquecidas para estudio de viroma ARN el tampón de lisis empleado no tenía adicionado el *carrier* ARN.

#### **4.5.2 Amplificación independiente de secuencia con un único oligonucleótido cebador (SISPA)**

El siguiente paso del flujo de trabajo realizado para el estudio de viroma (Figura 7. Página 88) fue el empleo del método SISPA en los extractos de las muestras enriquecidas para viroma ADN ó ARN

El estudio de viroma ADN y ARN no se diferencia únicamente en los tratamientos enzimáticos durante el enriquecimiento por tres métodos, además las PCRs empleadas en SISPA fueron también diferentes según el tipo de viroma a estudiar.



#### 4.5.2.1 SISPA para viroma ADN

Con el fin de amplificar el material genético viral enriquecido para el análisis de viroma ADN se realizó una amplificación no selectiva de ácidos nucleicos. Como control del proceso y como control negativo se empleó una muestra de agua libre de ADNasas estéril (Materiales. Apartado 3.5.1).

El método SISPA ADN empleado consistió en los siguientes pasos:

1) Síntesis de la primera hebra de ADN: partiendo de 8 µl de extracto enriquecido se realizó una incubación inicial con el oligonucleótido aleatorio FR26RV-N (Materiales. Tabla 12. Página 77) durante 2 minutos a 94°C con un enfriamiento posterior a 4°C. Tras esta incubación fueron añadidos los dNTPs a 10 mM, el fragmento de enzima *Klenow exopolimaerasa 3' → 5'* y su tampón correspondiente (Materiales. Apartado 3.5.1). El fragmento de *Klenow* es un fragmento proteico de gran tamaño perteneciente a la enzima ADN polimerasa I, procedente de *Escherichia coli*. Este fragmento mantiene la actividad polimerasa 3' → 5' que es usada para la síntesis de ADN de doble cadena a partir de muestras de cadena simple. La síntesis de doble cadena por la exopolimerasa *Klenow* ocurre durante 1 hora a 37°C con una posterior inactivación de la enzima durante 2 minutos a 94°C.

2) Síntesis de la segunda hebra de ADN de doble cadena: Tras la síntesis con la enzima *Klenow exopolimaerasa 3'→5'* inicial se realizó la segunda síntesis durante 1 hora a 37°C y la inactivación de la enzima a 75°C durante 20 minutos. El ADN obtenido fue purificado utilizando el kit *QIAquick PCR purification* de Qiagen (Materiales. Apartado 3.5.2).

3) Amplificación del ADN de doble cadena: Aprovechando la incorporación al azar del oligonucleótido FR26RV-N fue utilizado el oligonucleótido FR20RV (Materiales. Tabla 12. Página 77) para su amplificación inespecífica. Para esta segunda amplificación aleatoria fue preparada una mezcla que contenía 25 µl del reactivo *MyTaq Red Mix* (Materiales. Apartado 3.5.2) y el oligonucleótido FR20RV a 0,5 µM. A la mezcla fue añadido 10 µl del purificado anterior y se realizó una amplificación que consistió en 38 ciclos de 30 segundos a 98°C, 10 segundos a 98°C, 20 segundos a 54°C, 45 segundos a 72°C con una elongación final de 72°C durante 5 minutos.

Tras esta amplificación fue realizado un gel de agarosa 1% teñido con *Gel Red* (Materiales. Apartado 3.5.1) con el fin de analizar el difuminado característico de este método que presenta productos de amplificación con tamaños moleculares desde 200 pb hasta más de 1000 pb (Figura 8).

En base al difuminado obtenido se realizó una segunda amplificación de entre 6 y 10 ciclos añadiendo 12,5 µl de 2xMyTaq Red (Materiales. Apartado 3.5.1) y el oligonucleótido FR20RV (Materiales. Tabla 12. Página 77) a la mezcla previa.

Tras esta segunda PCR, la muestra para el estudio de viroma ADN estaba lista para la preparación de librerías para NGS.

#### 4.5.2.2 SISPA para viroma ARN

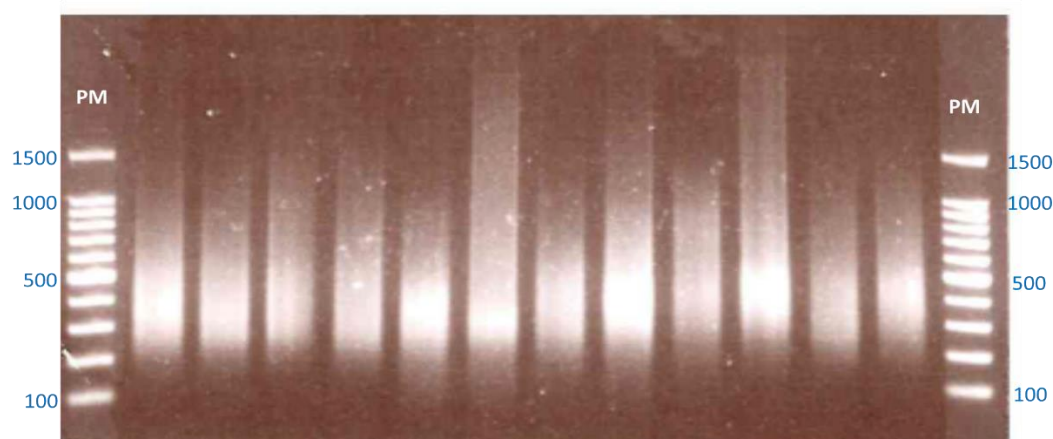
El procedimiento para realización del método SISPA para viroma ARN es similar al de viroma ADN con la diferencia de la necesidad de integrar el paso de transcripción inversa para transformar el ARN en ADNc en la síntesis de la primera hebra.

El método SISPA ARN empleado consistió en los siguientes pasos:

1) Síntesis de la primera hebra de ADNc: En este caso de nuevo se comenzó con una incubación de FR26RV-N (Materiales. Tabla 12. Página 77) y 100% dimetilsulfóxido (Materiales. Apartado 3.5.1) con 8µl del extracto enriquecido durante 5 minutos a 65°C con un enfriamiento posterior de la mezcla a 4°C durante 2 minutos. Para la obtención de ADNc fue usado el kit *SuperScript III First Strand Synthesis System* (Materiales. Apartado 3.5.2). A la mezcla anterior del extracto y el oligonucleótido fueron añadidos los siguientes reactivos aportados por el kit: ditiotreitól 100mM, Cloruro de Magnesio 25mM, dNTPS 10mM enzima ARNasa OUT para inhibir rinonucleasas como la ARNasa A y evitar la degradación del ARN, la enzima *SuperScript III* para la transcripción inversa del ARN y su tampón de uso provistos en el kit.

2) Síntesis de la segunda hebra de ADN complementario ADNc y eliminación del ARN molde: Una vez sintetizado el ADNc para la síntesis de la doble cadena fue añadida la enzima *Klenow exopolimaerasa 3'→5'* con la enzima ARNasa H provista por el kit *SuperScript III First Strand Synthesis System* (Materiales. Apartado 3.5.2) que degrada el ARN hibridado con ADN para obtener únicamente ADNc. La mezcla fue incubada durante 1 hora a 37°C y las enzimas inactivadas a 75°C durante 15 minutos.

El ADNc obtenido fue purificado utilizando el kit *QIAquick PCR purification* (Materiales. Apartado 3.5.2).



**Figura 8:** Imagen del degradado característico de SISPA

El protocolo de amplificación del ADNc de doble cadena es igual que el descrito para la amplificación de viroma ADN. Tras esta amplificación se visualizaron los productos mediante un gel de agarosa al 1% teñido con *Gel Red* (Materiales. Apartado 3.5.1) presenta productos de amplificación con tamaños moleculares desde 200 pb hasta más de 1000 pb (Figura 8).

Al igual que en el método para viroma ADN, en función del difuminado obtenido se realizó una segunda amplificación de entre 6 y 10 ciclos añadiendo 12,5 µl de 2xMyTaq Red y el oligonucleótido FR20RV a la mezcla previa.

Tras esta segunda PCR la muestra para el estudio de viroma ARN estaba lista para la preparación de librerías para NGS.

## 4.6. Amplificación por círculo rodante (RCA)

El método RCA fue el método de amplificación empleado para el estudio de virus circulares. Los reactivos empleados para esta amplificación procedían del kit *Illustra TempliPhi Amplification* (Materiales. Apartado 3.5.2).

Los ácidos nucleicos totales de interés fueron previamente extraídos por precipitación alcohólica (Métodos Apartado 4.4.1) para evitar contaminación por componentes ambientales similares a los virus circulares que pueden estar contenidas en las membranas de las columnas.

El protocolo consistió en la adición de 1 µl de extracto a 5 µl del tampón de muestra procedente del kit *Illustra TempliPhi Amplification*, esta mezcla fue incubada durante 3 minutos a 95°C.

A continuación se añadieron 5 µl del tampón de reacción y 0.2 µl de la enzima  $\phi$ 29 ADN polimerasa para proceder a una incubación a 30°C durante 4-18 horas. Tras la incubación, la ADN polimerasa fue inactivada a 65°C durante 10 minutos.

Con el fin de comprobar el rendimiento de la amplificación del virus circular de interés, se realizó un protocolo de linearización de los virus circulares empleando la enzima *NcoI* (Materiales. Apartado 3.5.1). El protocolo utilizado para la linearización consistió en una digestión enzimática durante 1 hora a 37°C del producto resultante del RCA con una posterior inactivación de la enzima a 67°C durante 10 minutos.

El producto de la linearización fue visualizado mediante electroforesis en gel de agarosa 1% teñido *Gel Red* (Materiales. Apartado 3.5.1) y se comprobó el tamaño de la banda obtenida.

## **4.7. Flujo de trabajo para la secuenciación de genoma completo del virus de la gripe**

Los protocolos de trabajo para WGS del virus de la gripe son diferentes a los establecidos para el estudio del viroma en pacientes pediátricos al no ser necesario ningún proceso de enriquecimiento viral.

Los ácidos nucleicos totales (Métodos. Apartado 4.4.2) se extrajeron empleando el kit *QIAamp Mini Elute Virus Spin* (Materiales. Apartado 3.5.2) en el extractor automático *Qiacube* (Materiales. Apartado 3.5.4). Después de la obtención del extracto fueron realizadas tres PCRs previamente publicadas para amplificación de genoma completo del virus de la gripe: 1. PCR multisegmento Gripe A [16], 2. PCR multisegmento Gripe B [17] y 3. PCR multisegmento universal [18].

El producto de estas amplificaciones fue visualizado mediante electroforesis en gel de agarosa 1% teñido con *Gel Red* (Materiales. Apartado 3.5.1).

### **4.7.1 PCR multisegmento Gripe A.**

Para la PCR de amplificación de los 8 segmentos de gripe A fue usado el kit *Superscript III One-step RT-PCR System con Platinum Taq High Fidelity* (Materiales. Apartado 3.5.2). Se preparó una mezcla de 25 µl compuesto por el tampón de reacción 2X y los oligonucleótidos 0,16 µM

Uni12/Inf1, 0,24  $\mu$ M del Uni12/Inf3 y 0,4  $\mu$ M Uni13/Inf1 (Materiales. Tabla 12. Página 77) a los que fueron añadidos 10 $\mu$ l del extracto de ARN. Las condiciones de amplificación de esta PCR fueron una retrotranscripción inicial de 1 hora a 42°C y una desnaturalización de 2 minutos a 94°C. A continuación 5 ciclos de amplificación de 30 segundos a 94°C, 30 segundos a 44°C y 3 minutos a 68°C. Tras estos 5 ciclos fueron realizados otros 35 ciclos de amplificación de 30 segundos a 94°C, 30 segundos a 58°C, 3 minutos a 68°C y una elongación final de 5 minutos a 68°C.

#### 4.7.2 PCR multisegmento Gripe B.

Para la amplificación de los ocho segmentos del genoma de gripe B fue seguido un protocolo de amplificación previamente publicado [17].

El kit empleado para la preparación de la mezcla de reacción fue el kit *Superscript III One-step RT-PCR System con Platinum Taq High Fidelity* (Materiales. Apartado 3.5.2). La mezcla de reacción estuvo compuesta por el tampón de reacción 2X RT con la enzima *Superscript III*, la mezcla de oligonucleótidos de gripe B (Materiales. Tabla 12. Página 77) y 5  $\mu$ l de extracto.

#### 4.7.3 PCR multisegmento universal de gripe

Para la amplificación del genoma completo de ambos tipos de gripe conjuntamente el programa de la amplificación empleado es el mismo programa que el de la PCR multisegmento de Gripe A (Métodos. Apartado 4.7.1). El kit empleado para la mezcla de reacción fue *Superscript III One-step RT-PCR System con Platinum Taq High Fidelity* (Materiales. Apartado 3.5.2) como en las PCRs multisegmento descritas anteriormente. El único punto en el que difiere este protocolo del de la PCR multisegmento gripe A fue el uso de 3 oligonucleótidos diferentes: uniFlu, uniFluF y uniFluR (Materiales. Tabla 12. Página 77).

### 4.8. Preparación de librerías Nextera XT

Para la preparación de librerías se empleó el protocolo recomendando del kit de preparación de librerías *Nextera XT* [162] (Materiales. Apartado 3.5.2).

La preparación de librerías siguiendo este método tiene 5 puntos fundamentales (Figura 9): 1. la tagmentación; 2. la amplificación de librerías; 3. purificación de librerías; 4. normalización de librerías y 5. la mezcla de librerías.

La tagmentación del método *Nextera XT* emplea un transposón que en un solo paso fragmenta el ADN y añade los adaptadores.

En el paso en el que ocurre la amplificación de librerías fueron añadidos los índices procedentes del kit de índices *Nextera XT v2* (Materiales. Apartado 3.5.2), estos índices son secuencias que a modo de etiqueta permiten distinguir entre las diferentes muestras en la mezcla que se produce previa la secuenciación y es un paso clave de este protocolo.

Tras la amplificación de las librerías fueron purificadas empleando el reactivo *Ampure XP* (Materiales. Apartado 3.5.1) este reactivo está formado por bolas magnéticas que se unen selectivamente a los fragmento de ADN según su concentración.

Uno de los puntos de control del proceso y de evaluación de la calidad de las librerías preparadas es el análisis de la distribución de tamaño de las librerías producidas empleando un chip de ADN de alta sensibilidad *Agilent* (Materiales. Apartado 3.5.1) en un bioanalizador *Agilent* (Materiales. Apartado 3.5.4). Los resultados obtenidos del análisis de tamaños y concentración de las librerías son empleados en la normalización de las librerías para obtener la misma representación de cada una de ellas en la mezcla. La mezcla de librerías fue cargada a una concentración de 1 ng en el secuenciador *MiSeq* (Materiales. Apartado 3.5.4) empleando el kit de reactivos *MiSeq v2* de 300 ciclos (Materiales. Apartado 3.5.2).



**Figura 9:** Flujo de trabajo del protocolo Nextera XT para la preparación de librerías.  
Figura adaptada de *Nextera XT Library Preparation Kit Reference Guide*.

## 4.9. Método de enriquecimiento *HyperCap Target Enrichment* con preparación de librerías con el kit *Kapa Hyper Plus*

Debido a la necesidad del empleo de un método de enriquecimiento alternativo al enriquecimiento por 3 métodos (Métodos. Apartado 4.5.1) fue usado el panel de captura por sondas *VirCapSeq* alojado en el kit *HyperCap Target Enrichment* (Materiales. Apartado 3.5.2).

Para la preparación de librerías fue seguido el protocolo recomendado por el kit *Kapa Hyper Plus*, elegido por ser este un método de preparación de librerías compatible con el método de enriquecimiento usado (Materiales. Apartado 3.5.2). En el protocolo empleado de preparación de librerías el proceso de fragmentación es enzimático y a los extremos se les adicionó una cola poli-A para la unión de los adaptadores e índices de los kits de índices *SeqCap Adapter A y B* (Materiales. Apartado 3.5.2).

Las librerías finalizadas fueron purificadas usando el reactivo *Ampure XP* (Materiales. Apartado 3.5.1) y analizadas usando un chip de ADN de alta sensibilidad *Agilent* (Materiales. Apartado 3.5.1) en un bioanalizador *Agilent* (Materiales. Apartado 3.5.1) para ver la calidad y tamaño de las librerías.

Una vez comprobada la calidad de las librerías se produjo la mezcla de éstas a una concentración de 1 µg.

El método de enriquecimiento se basó en la incubación durante 20 horas de la mezcla de librerías junto con la mezcla de sondas procedentes del kit *HyperCap Target Enrichment*.

La mezcla de librerías enriquecidas fue cargada a una concentración de 1 ng en el secuenciador *MiSeq* (Materiales. Apartado 3.5.4) empleando el kit de reactivos *MiSeq v2* de 300 ciclos (Materiales. Apartado 3.5.2).



**Figura 10:** Flujo de trabajo del protocolo de preparación de librerías *Kapa Hyper Plus* y de enriquecimiento con el panel de captura por sondas *HyperCap Target Enrichment*

## 4.10. Análisis bioinformático de las secuencias obtenidas por NGS

### 4.10.1 Análisis bioinformático de las secuencias procedentes del estudio de viroma y de adenovirus de murciélagos

El análisis bioinformático de las secuencias obtenidas por NGS en el estudio de viroma y de adenovirus de murciélagos se realizó empleando dos plataformas de análisis diferentes: la plataforma *Virus TAP* [24] y una línea de códigos para el ensamblaje desarrollada y mantenida por el Centro de Ciencias Genómicas USAMRIID (Fort Detrick, Frederick, Maryland, USA).

La plataforma *Virus TAP* [24] es una plataforma web de fácil uso en la que se alojan diferentes herramientas para el análisis de secuencias. El flujo de análisis de esta plataforma comienza con la selección de las secuencias que pasan el filtro de calidad empleando la herramienta *Skewer* [163] (Materiales. Apartado 3.5.5). Posteriormente la eliminación de secuencias de genomas no virales fue realizada con el programa *Bwasw v.0.7.9a-r786*, (Materiales. Apartado 3.5.5), este



programa detectó y eliminó las secuencias de ARN ribosomal, genoma bacteriano y genoma del hospedador.

Un segundo análisis para descartar la secuencias no virales fue realizado ensamblando las secuencias disponibles tras el primer filtrado con la herramienta *IDBA-UD* [164], y analizando dichas secuencias con los programas *Megablast*, v.2.2.26, [165] y *RAP Search2* (v. 2.16) [166] (Materiales. Apartado 3.5.5) comparándolas contra las secuencias de nucleótidos/proteínas virus presentes las bases de datos de *GenBank* (Materiales. Apartado 3.5.6). Los valores no significativos de este proceso se determinaron como fragmentos no virales y fueron eliminados.

Tras la eliminación de los fragmentos no virales se produjo el ensamblaje *de novo* de las secuencias. La plataforma *Virus TAP* aloja 4 líneas de código diferentes para el ensamblaje: *A5-miseq* [167], *Platanus*, version1.2.1, [168] con *PriceTI* [169], y *IDBA-UD* [164] con *PriceTI* [169] (Materiales. Apartado 3.5.5). En esta tesis fue empleado el ensamblaje empleando la línea de código *A5-miseq* [167] (Materiales. Apartado 3.5.5).

El análisis de datos realizado en USAMRIID empleó herramientas de análisis bioinformático diferentes a las empleadas por la plataforma *Virus TAP* [24]. La selección de secuencias de calidad y la eliminación de adaptadores, oligonucleótidos y duplicados de PCR fue realizado empleando las herramientas *Cutadapt* v1.7 [170], *Prinseq-lite* v.0.20.4 [171] y *Picard* (<http://broadinstitute.github.io/picard>) (Materiales. Apartado 3.5.5). Para la eliminación del genoma del hospedador se empleó la herramienta *Bowtie2* v.2.1.6 [172], (Materiales. Apartado 3.5.5) que alineó las secuencias que superaron el filtrado de calidad frente al genoma humano de referencia GRCh38 (<https://www.ncbi.nlm.nih.gov/genome/guide/human>). El ensamblaje de las secuencias restantes fue realizado por síntesis *de novo* con la herramienta *Ray* v.2.2, [173] (Materiales. Apartado 3.5.5). Los fragmentos ensamblados fueron identificados taxonómicamente a través de la similitud de las secuencias con otras presentes en la base de datos *GenBank* usando *BLAST* v2.2.28+ (megablast y dc-megablast) (Materiales. Apartado 3.5.6).

Las lecturas que no fueron asignadas taxonómicamente se alinearon con las secuencias almacenadas en la base de datos *GenBank* empleando el programa *DIAMOND* [174] en modo sensible para confirmar que esas lecturas ni se asocian a virus ni son potenciales nuevos patógenos.

#### 4.10.2 Análisis bioinformático de las muestras de gripe.

Para el análisis bioinformático de las secuencias de los virus de la gripe fueron usadas las plataformas *IRMA* [21] e *INSaFLU* [25] específicas para gripe. La plataforma de análisis de gripe *INSaFLU* es una plataforma web [25] que genera de forma automática las secuencias de los ocho segmentos de gripe procedentes de NGS.

El control de calidad de las secuencias obtenidas y la eliminación de lecturas y bases de baja calidad, de índices y adaptadores se realizó mediante el empleo del programa *FastQC* para el control de calidad de las secuencias, <https://www.bioinformatics.babraham.ac.uk/projects/fastqc>, y la herramienta *Trimmomatic* para la eliminación de los elementos previamente descritos [175].

Para la identificación de los tipos y los subtipos de gripe se realizó el ensamblaje de novo de las secuencias que pasaron el control de calidad usando el ensamblador *SPAdes* [176]. Posteriormente, la herramienta *ABRicate*, <https://github.com/tseemann/abricate>, fue empleada para comparar las secuencias disponibles con la base de datos que contiene la información de los genes marcadores que permiten la discriminación entre gripe A y B y entre los subtipos de gripe A y los dos linajes existentes de gripe B.

Para la generación de las secuencias consenso y la detección de variantes se empleó el multisoftware *Snippy*, <https://github.com/tseemann/snippy>, esta herramienta fue empleada en el alineamiento de secuencias, el análisis de la aparición de inserciones-deleciones, la anotación de variantes y la generación de consensos empleando una secuencia de referencia aportada por el usuario.

El meta-ensamblador *IRMA* [21] alojado en la plataforma *Galaxy*, <https://usegalaxy.org/>, para facilitar su manejo al estar en un entorno web, generó las secuencias consenso y las subpoblaciones virales de los virus de la gripe secuenciados. Además de las secuencias los datos generados fueron acompañados de gráficos representativos de las variantes minoritarias, la cobertura de los segmentos del genoma del virus y la posibilidad de la existencia en un mismo segmento de las diferentes variantes minoritarias.

En este caso el control de calidad de las secuencias se realizó también empleando el software *FastQC*, <https://www.bioinformatics.babraham.ac.uk/projects/fastqc>, y la herramienta *Trimmomatic* [175].

Tras la eliminación de lecturas y bases de baja calidad, de índices y adaptadores se generaron las secuencias de los segmentos de gripe por refinamiento iterativo empleando la herramienta *LABEL* que agrupa las lecturas y va modificando las secuencias de referencia teniendo en cuenta la alta diversidad poblacional y las tasas de mutación [177] .

#### 4.11. Diseño de oligonucleótidos

Para la realización del diseño de los oligonucleótidos cebadores las secuencias de interés y de referencia depositadas en la base de datos *GenBank* (Materiales. Apartado 3.5.6) se alinearon de manera múltiple utilizando el programa *MACAW*, v2.05, (Materiales. Apartado 3.5.5).

Tras el alineamiento de las secuencias se definieron zonas con una homología que les hizo ser candidatas a definir la secuencia concreta de un determinado oligonucleótido. Todos los oligonucleótidos presentes en la Tabla 13 fueron diseñados siguiendo esta metodología.

#### 4.12. PCRs tiempo real

##### 4.12.1 PCR múltiples tiempo real BRQ y PCR múltiple tiempo real respiratorios 2

El empleo de estas PCRs tiempo real engloba el diagnóstico de los principales virus respiratorios, excluyendo al virus de la gripe, y son empleadas en todas las muestras respiratorias pediátricas recibidas de los proyectos PI15CIII/00028 y PI18CIII/00009.

La PCR múltiple tiempo real BRQ es empleada para la detección de VRS, HMNV, AdV y BOV. La PCR múltiple tiempo real respiratorios 2 es empleada para el diagnóstico de los 4 tipos de VPI y el grupo de virus perteneciente al género Enterovirus. Los oligonucleótidos y sondas empleados en estas PCRs están descritos en las Tablas 13 y 15.

En la mezcla de reacción de esta PCR fueron añadidos 5 µL de los ácidos nucleicos totales extraídos a 20 µL de una mezcla que contenía los oligonucleótidos descritos en la Tabla 12 y las correspondientes sondas descritas en la Tabla 14, página 78, sondas específicas para cada virus, el tampón de la polimerasa 2x y la polimerasa *SuperScript III* procedentes del kit *SuperScript III PlatinumOne-Step RT-PCR System* (Materiales. Apartado 3.5.2).

Las condiciones de amplificación fueron una RT inicial a 52°C durante 30 minutos y una desnaturalización de 2 minutos a 95°C. Tras la RT y la desnaturalización se produjeron 50 ciclos de 95°C durante 15 minutos y 55°C durante 1 minuto y 30 segundos.

Los resultados del ensayo en cada canal de fluorescencia se consideraron positivos si la fluorescencia observada durante la amplificación era superior a 100 unidades relativas de fluorescencia (RFUs) y seguía una curva sigmoidea.

En los resultados positivos fue anotado, además, el número de ciclos en que el valor de fluorescencia superó el límite establecido (100 RFUs) denominado umbral de ciclo o Ct.

Los perfiles de amplificación en el ensayo con los diferentes fluoróforos en el caso de la PCR múltiple BRQ fueron VRS con el fluoróforo FAM, Adv con VIC, HMNV con Cy5 y BoV con TYE705.

En el caso de la PCR múltiple respiratorios 2 los perfiles de amplificación en el ensayo con los diferentes fluoróforos fueron VPI con el fluoróforo FAM y los EV con HEX.

#### **4.12.2 La PCR de identificación de Gripe tipos A, B y C**

La PCR de identificación de los virus de la gripe A, B y C es el método empleado en el Laboratorio de Virus Respiratorios y Gripe del CNM para la caracterización de los virus de la gripe.

El protocolo está acreditado en calidad según Norma ISO 15189 (VRP PT 006 03) desarrollado e implementado en el Laboratorio.

Los oligonucleótidos y las sondas empleadas están descritos en las Tablas 13 y 15.

La mezcla de reacción de esta PCR constó de 5 µL de los ácidos nucleicos totales extraídos añadidos a 20 µL de una mezcla que contenía los oligonucleótidos descritos en la Tabla 12 y las correspondientes sondas descritas en la Tabla 14, página 78, el tampón 2x de la polimerasa y la polimerasa *SuperScript III* procedentes del kit *SuperScript III PlatinumOne-Step RT-PCR System* (Materiales. Apartado 3.5.2).

El programa de amplificación consistió en una RT inicial a 50°C durante 30 minutos con la posterior desnaturalización de 2 minutos a 95°C y 45 ciclos de 95°C durante 15 segundos y 55°C durante 1 minuto.

Las condiciones para considerar el resultado positivo fueron las mismas descritas previamente (Métodos. Apartado 4.12.1).

Los perfiles de amplificación en el ensayo con los diferentes fluoróforos fueron gripe A con el fluoróforo FAM, gripe B con VIC y gripe C con Cy5.

#### 4.12.3 PCR MERS-CoV

La PCR MERS-CoV es el método de diagnóstico para este virus y es el método recomendado por la OMS. El protocolo está acreditado en calidad según Norma ISO 18535 (VRP PT 008 03) desarrollado e implementado en el Laboratorio.

Los oligonucleótidos y sondas empleadas están descritas en las Tablas 13 y 15.

En la mezcla de reacción de esta PCR fueron añadidos 5 µL de los ácidos nucleicos totales extraídos a 20 µL de una mezcla que contenía los oligonucleótidos descritos en la Tabla 12 y las correspondientes sondas descritas en la Tabla 14, página 78, específicos para cada PCR, el tampón de la polimerasa 2x y la polimerasa *SuperScript III* procedentes del kit *SuperScript III PlatinumOne-Step RT-PCR System* (Materiales. Apartado 3.5.2).

El protocolo de amplificación empleado fue una RT inicial a 55°C durante 20 minutos seguida de una desnaturalización a 94°C durante 3 minutos, posteriormente se produjeron 45 ciclos de 94°C durante 15 segundos y 58°C durante 30 segundos.

Las condiciones para considerar el resultado positivo fueron las mismas descritas previamente (Métodos. Apartado 4.12.1). El fluoróforo empleado en este ensayo fue FAM para ambos genes.

### 4.13. PCR convencionales

#### 4.13.1 PCR de estudio REP HRCiV

La PCR HRCiV 2 fue empleada en el estudio del virus circular HRCiV.

Los oligonucleótidos empleados fueron diseñados durante el desarrollo de la tesis y están descritos en la Tabla 13, página 78.

En la mezcla de reacción de esta PCR fueron añadidos 5 µL de los ácidos nucleicos totales extraídos a 20 µL de una mezcla que contenía los oligonucleótidos descritos en la Tabla 13 (HRCiV2F y HRCiV2R), el tampón de la polimerasa 2x y la polimerasa procedentes del kit *Lightcycler 480 probes master* (Materiales. Apartado 3.5.2).

El programa de amplificación consistió en una desnaturalización a 95°C durante 5 minutos previa a 45 ciclos de 30 segundos a 95°C y 1' 30" a 64°C.

#### 4.13.2 PCR de genoma completo de HRCiV

Con el fin de amplificar el genoma completo del virus HRCiV para obtener su secuencia por el método Sanger fueron diseñadas dos PCRs: genoma completo HRCiV 1 y 2.

#### 4.13.2.1 PCR de genoma completo 1

Para la preparación de la mezcla de reacción de esta PCR fueron añadidos 5 µL de los ácidos nucleicos totales extraídos a 20 µL de una mezcla que contenía los oligonucleótidos descritos en la Tabla 13 (HRCiVT1 y HRCiVT4), el tampón 10x Taq polimerasa con  $(\text{NH}_4)_2\text{SO}_4$  y  $\text{MgCl}_2$  20mM, la mezcla de 0,2 mM de cada dNTP y 2.5 unidades de la polimerasa *AmpliTaq* (Materiales. Apartado 3.5.1).

El protocolo de amplificación consistió en una desnaturalización a 95°C durante 5 minutos, seguida de 45 ciclos de 30 segundos a 95°C, 1 minuto a 60°C y 2 minutos a 72°C.

#### 4.13.2.2 PCR de genoma completo 2

La mezcla de reacción de esta PCR consistió en 5 µL de los ácidos nucleicos totales extraídos añadidos a 20 µL de una mezcla que contenía los oligonucleótidos descritos en la Tabla 13, página 78, (HRCiVT5 y HRCiVT8), el tampón 10x Taq polimerasa con  $(\text{NH}_4)_2\text{SO}_4$  y  $\text{MgCl}_2$  20mM, la mezcla de 0,2 mM de cada dNTP y 2.5 unidades de la polimerasa *AmpliTaq* (Materiales. Apartado 3.5.1). El programa de amplificación consistió en una desnaturalización a 95°C durante 5 minutos, tras esto 45 ciclos de 30 segundos a 95°C, 1 minuto a 57°C y 2 minutos a 72°C.

#### 4.13.3 PCR de hexón para adenovirus

Con el fin de estudiar la presencia de AdV en murciélagos se usó una PCR anidada diseñada para amplificar una de las siete regiones hipervariables del gen del hexón.

En la primera PCR fueron añadidos 5 µL de los ácidos nucleicos totales extraídos en las muestras respiratorias y de heces a 45 µL de una mezcla que contenía a los oligonucleótidos descritos en la Tabla 12 (gADV1s y gADV1r) a una concentración de 1,2 µM, el tampón 10x Taq polimerasa con  $(\text{NH}_4)_2\text{SO}_4$  y  $\text{MgCl}_2$  20mM, la mezcla de dNTPs a 0,2 mM y 2.5 unidades de la polimerasa (Materiales. Apartado 3.5.1).

La condiciones de amplificación para esta mezcla fueron una activación de 95°C durante 4 minutos y una amplificación de 40 ciclos de 95°C durante 30 segundos, 50°C durante 2 minutos y 72°C durante 30 segundos.

Para la segunda reacción de amplificación anterior fueron añadidos 2µL, del producto de la PCR previa, a una mezcla de 48 µL que contenía los reactivos mencionados anteriormente siendo los oligonucleótidos HADV2+ y panADV1R a 1,2µM los utilizados en la segunda reacción (Tabla 12.

Página 77). Las condiciones fueron similares a la primera ronda produciéndose un único cambio empleando 49°C como temperatura de alineamiento.

#### 4.13.4 PCR ADN polimerasa para adenovirus

Para alcanzar el objetivo de comparar nuestras muestras positivas con trabajos previos y mejorar la exactitud de nuestras filogenias se desarrolló una PCR semianidada que amplificaba un fragmento del gen de la ADN polimerasa ADN dependiente.

En la primera amplificación de este fragmento se añadieron 5 µl de los ácidos nucleicos extraídos a 20 µl de una mezcla de reacción que contenía los oligonucleótidos polF y polR2 a una concentración de 0,4 µM (Tabla 13. Página 78), el tampón 2x de la polimerasa y la polimerasa procedentes del kit *Lightcycler 480 probes master* (Materiales. Apartado 3.5.2).

Las condiciones de amplificación fueron una activación de 95°C durante 5 minutos seguido de la amplificación de 45 ciclos de 95°C durante 30 segundos, 57°C durante 2 minutos y 68°C durante 30 segundos.

Para la segunda reacción de amplificación fueron añadidos 2 µl del producto de la amplificación previa a 23 µl de una mezcla que contenía los oligonucleótidos polF y polR2 a una concentración de 0,4 µM (Tabla 13. Página 78), tampón 10x *Taq* polimerasa con (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> y MgCl<sub>2</sub> 20mM, la mezcla de dNTPs a 0,2 mM y 1.25 unidades de la polimerasa *AmpliTaq* (Materiales. Apartado 3.5.1). Las condiciones de amplificación fueron 95°C durante 5 min de activación seguido de una amplificación de dos pasos de 95°C durante 15 segundos y 62°C durante 2 minutos en 45 ciclos.

#### 4.13.5 Visualización de productos de amplificación de las PCRs convencionales.

El producto de amplificación fue visualizado por electroforesis en geles de agarosa al 2% teñidos con *GelRed* (Materiales. Apartado 3.5.1).

### 4.14. Purificación del ADN y secuenciación

Con el fin de obtener las secuencias de las muestras positivas detectadas por PCRs convencionales los productos de amplificación fueron purificados y secuenciados.

El ADN amplificado fue purificado con el kit *QIAquick PCR purification de Qiagen* (Materiales. Apartado 3.5.2). Tras la purificación fueron secuenciados bidireccionalmente usando el kit *Big Dye Terminator v.3.1* (Materiales. Apartado 3.5.2. Página 78) en un secuenciador automático *ABI*

*PRISM 3700 DNA Analyzer* (Materiales. Apartado 3.5.4). El análisis, ensamblaje y obtención de las secuencias consenso fue realizado usando los programas bioinformáticos *Chromas PRO v1.5* y *EditSeq* (Materiales. Apartado 3.5.6).

## 4.15. Análisis filogenéticos

En los objetivos II, III y IV fueron realizados análisis filogenéticos a partir de las secuencias obtenidas por NGS o por método Sanger.

Las secuencias consenso obtenidas fueron comparadas con secuencias publicadas y depositadas en las bases de datos *GenBank* y *GISAID* (Materiales. Apartado 3.5.6) usando el algoritmo *BLASTn* (Materiales. Apartado 3.5.5) con el fin encontrar las secuencias publicadas en las bases de datos que pudieran ser de interés para el análisis.

Para el análisis las secuencias obtenidas fueron alineadas de manera múltiple junto con la selección de secuencias obtenidas de las bases de datos previamente descritas usando el programa *CLUSTAL X v2.0* (Materiales. Apartado 3.5.5).

El análisis de distancias entre secuencias y el análisis filogenético se realizó empleando modelos filogenéticos seleccionados para cada grupo de virus a analizar mediante el programa *Modeltest* [178] (Materiales. Apartado 3.5.5).

Para el cálculo del respaldo estadístico que definen las relaciones filogenéticas y la representación gráfica de estos cálculos mediante árboles filogenéticos se utilizó el programa *MEGA 5.2* (Materiales. Apartado 3.5.5).

### 4.15.1 Análisis filogenético de rinovirus detectados en pacientes pediátricos

El análisis filogenético de rinovirus de referencia y los rinovirus detectados en esta tesis y descritos en el Apartado 5.2.1.3 del Objetivo II fue realizado con el programa *MEGA 5.2* usando el método *Neighbor-Joining* con el modelo *Tamura Nei* con un respaldo estadístico de 5000 repeticiones para el análisis de la secuencia de nucleótidos de estos virus.

La identificación de *GenBank* (Materiales. Apartado 3.5.6) de las secuencias empleadas fue: “Rinovirus humano A21 isolate rmh123/2013” (KR090587), “Rinovirus humano 55” (DQ473511), “Rinovirus humano 81 strain ATCC VR-1191” (FJ445157), “Rinovirus humano 1A strain HRV-1A\_P25” (KC894169), “Rinovirus humano 43 strain ATCC VR-1153” (FJ445131), Rinovirus humano 64 strain ATCC VR-1174 (FJ445181), “Rinovirus humano 15” (DQ473493), “Rinovirus humano 22 strain ATCC VR-1132” (FJ445122), “Rinovirus humano 13 strain ATCC VR-1123” (FJ445116),



“Rinovirus humano 67 strain ATCC VR-1177” (FJ445149), “Rinovirus humano 9 isolate F01” (FJ445114), “Rinovirus humano 2” (X02316), “Rinovirus humano 8 strain ATCC VR-1118” (FJ445113), “Rinovirus humano 20 strain ATCC VR-1130” (FJ445120), “Rinovirus humano 30 strain ATCC VR-1140” (FJ445179) y “Rinovirus humano 45 strain ATCC VR-1155” (FJ445132).

#### **4.15.2 Análisis filogenético de la familia Redondoviridae**

El análisis filogenético de los virus pertenecientes a la familia *Redondoviridae* y los virus HRCiV caracterizados en esta tesis representado en el Apartado 5.3.2.5 del Objetivo II fue realizado con el programa *MEGA 5.2* usando el método *Maximum Likelihood* con el modelo *Hasegawa-Kishino-Yano* con un respaldo estadístico de 5000 repeticiones para el análisis de la secuencia de nucleótidos de estos virus.

La identificación de *GenBank* (Materiales. Apartado 3.5.6) de las secuencias empleadas fue: “Porcine stool-associated circular virus 5 isolate CP3” (PosCV5, KJ433989), “human respiratory-associated PSCV-5-like virus” (HRAPLV, KY052047), “Human gut-associated brisavirus VW” (MK059759), “Human oral-associated videntovirus XM, MC, LZ, EC, AV” (MK059771, MK059770, MK059769, MK059768, MK059767), “Human lung-associated videntovirus FB, LT, JY, JB, ES, DC, AL” (MK059763, MK059766, MK059765, MK059764, MK059762, MK059761, MK059760), “Human oral-associated brisavirus YH” (MK059758), “Human lung-associated brisavirus RC, MD” (MK059757, MK059756), “Human lung-associated brisavirus II” (MK059755) y “Human lung-associated brisavirus AA” (MK059754).

#### **4.15.3 Análisis filogenético de los virus de la gripe B linaje Yamagata de los pacientes pediátricos con cuadro gripal**

En el caso del estudio de los segmentos del virus gripe B linaje Yamagata de interés se realizaron dos análisis filogenéticos empleando los mismos parámetros que fueron representados en el Apartado 5.3.5 del Objetivo III. Estos análisis se realizaron con el programa *MEGA 5.2* usando el método *Maximum Likelihood* con el modelo *Hasegawa-Kishino-Yano* con un respaldo estadístico de 5000 repeticiones.

El primer análisis realizado empleó el segmento de la matriz de virus de la gripe tipo B cuyas secuencias fueron obtenidas de la base de datos *GISAID* (Materiales. Apartado 3.5.6) para el análisis de la secuencia de nucleótidos de estos virus. El segundo análisis se realizó empleando los ocho segmentos de los virus de la gripe obtenidos en la base de datos *GISAID* junto con los segmentos de los virus de la gripe B linaje Yamagata de los pacientes pediátricos 1 y 2.

Las secuencias empleadas fueron: “B/Norway/2409/2017” (EPI\_ISL\_290858), “B/Maine/14/2018” (EPI\_ISL\_306182), “B/Brisbane/3/2007” (EPI\_ISL\_149491), “B/Florida/4/2006” (EPI\_ISL\_22808), “B/Phuket/3073/2013” (EPI\_ISL\_336681) y “B/West Virginia/28/2018” (EPI\_ISL\_313174).

#### **4.15.4 Análisis filogenético del fragmento del hexón de adenovirus.**

En el caso del análisis filogenético en el estudio del fragmento de hexón de AdV descrito en el apartado 5.4.2 el método empleado fue *Neighbor-Joining* con el modelo *Tamura 3 parámetros* usando un test de filogenia con respaldo estadístico de 5000 repeticiones.

Las secuencias utilizadas para el análisis filogenético obtenidas de la base de datos *GenBank* (Materiales. Apartado 3.5.6) fueron las siguientes: BtAdV-3 (cepa TJM, GU226970, mastadenovirus A de murciélago), BtAdV-2 (cepa PPV1, JN252129, mastadenovirus B de murciélago), AdVs humanos: tipo 1 (AF534906), tipo 2 (J01917), tipo 3 (DQ086466), tipo 4 (AY594254), D8 cepa Ger/Berlin/04\_2003 (KT862545), tipo 9 (AJ854486), tipo 12 (X73487), tipo 14 (FJ841902), tipo 16 (X74662), tipo 21 (KF528688), tipo 24 (JN226751), tipo 27 (JN226753), tipo 42 (JN226761), tipo 45 (JN226764); AdVs de simio : tipo 1 (AY771780), tipo 4 (KP853121); AdVs ovinos: tipo 1 (DQ630754), tipo 3 (DQ630756); AdV 5 porcino (AF289262); AdVs murino: tipo 1 (M81889), tipo 3 (EU835513); AdVs bovino: tipo 1 (DQ630761), tipo 10 (AF282774); AdVs canino: tipo 1 (KX545420), tipo 2 (U77082), AdV equipo tipo 2 (L80007), mastadenovirus de murciélagos: BtAdV-7 (cepa WIV12, KT698856, mastadenovirus D de murciélagos), BtAdV-8 (cepa WIV13, KT698852, mastadenovirus E de murciélagos), BtAdV-9 (cepa WIV17, KX961095, mastadenovirus F de murciélagos), *Rousettus leschenaultii* WIV18 (NC\_035072), BtAdV-4 (cepa WIV9, KT698853, mastadenovirus C de murciélagos), *Rhinolophus sinicus* WIV10 (NC\_029899), *R. sinicus* WIV11 (NC\_029902), BtAdV-11 (cepa 250-A, KX871230, mastadenovirus G de murciélagos).

#### **4.15.5 Análisis filogenético del fragmento de ADN polimerasa de adenovirus**

En el caso del análisis filogenético en el estudio del fragmento de ADN polimerasa de AdV descrito en el apartado 5.4.3 el método empleado fue *Neighbor-Joining* con el modelo *Kimura 2 parámetros* usando un test de filogenia con respaldo estadístico de 5000 repeticiones.

En este caso la relación de secuencias utilizadas obtenidas de la base de datos *GenBank* (Materiales. Apartado 3.5.6) fueron: BtAdV-3 (cepa TJM, GU226970, mastadenovirus A de murciélago), BtAdV-2 (cepa PPV1, JN252129, mastadenovirus B de murciélago), human AdVs:

tipo 1 (AF534906), tipo 2 (J01917), tipo 3 (DQ086466), tipo 4 (AY594254), tipo 5 (AY339865), tipo 7 (AY594256), tipo 6 (HQ413315), tipo 9 (AJ854486), tipo 12 (X73487), tipo 17 (AF108105), tipo 19 (JQ326209), tipo 26 (EF153474), tipo 48 (EF153473), tipo 53 (AB605245); AdVs de simio: tipo 1 (AY771780), tipo 4 (KP853121); AdVs bovinos: tipo 2 (AF252854), tipo 3 (AF061654), tipo 1 (AC\_000191); AdVs porcino: tipo 3 (AB026117), tipo 5 (AF289262); AdVs canino: tipo 1 (KX545420), tipo 2 (U77082); mastadenovirus de murciélagos: BtAdV-7 (cepa WIV12, KT698856, mastadenovirus D de murciélago), BtAdV-8 (cepa WIV13, KT698852, mastadenovirus E de murciélago), BtAdV-9 (cepa WIV17, KX961095, mastadenovirus F de murciélago), *Rousettus leschenaultii* WIV18 (NC\_035072), BtAdV-4 (cepa WIV9, KT698853, mastadenovirus C de murciélago), *Rhinolophus sinicus* WIV10 (NC\_029899), *R. sinicus* WIV11 (NC\_029902), BtAdV-11 (cepa 250-A, KX871230, mastadenovirus G de murciélago), *Plecotus austriacus* Kupa (JN167523), *Rhinolophus ferrumequinum* Teresztenye (JN167522), *Myotis* spp 1419 (GU226962), *R. leschenaultii* 1050597 (HQ529709), *Nyctalus noctula* 2 119/08 (KM043096), *Myotis emarginatus* BS18 (KM043084), *Myotis myotis* BS16 (KM043106), *Myotis blythii* GTHU38 (KM043086), *Nyctalus leisleri* 239/08 (KM043102), *Pipistrellus pygmaeus* GTHU64 (KM043090), *P. pygmaeus* 176/09 (KM043091).

# Resultados y Discusión



## 5.1 OBJETIVO I: Desarrollo del protocolo de trabajo para el estudio de viroma en muestras clínicas respiratorias

La utilización de métodos NGS en los laboratorios de referencia puede aportar importantes ventajas al diagnóstico de las IRA de origen viral aunque su aplicación conlleva un exhaustivo trabajo de puesta a punto y estandarización de protocolos.

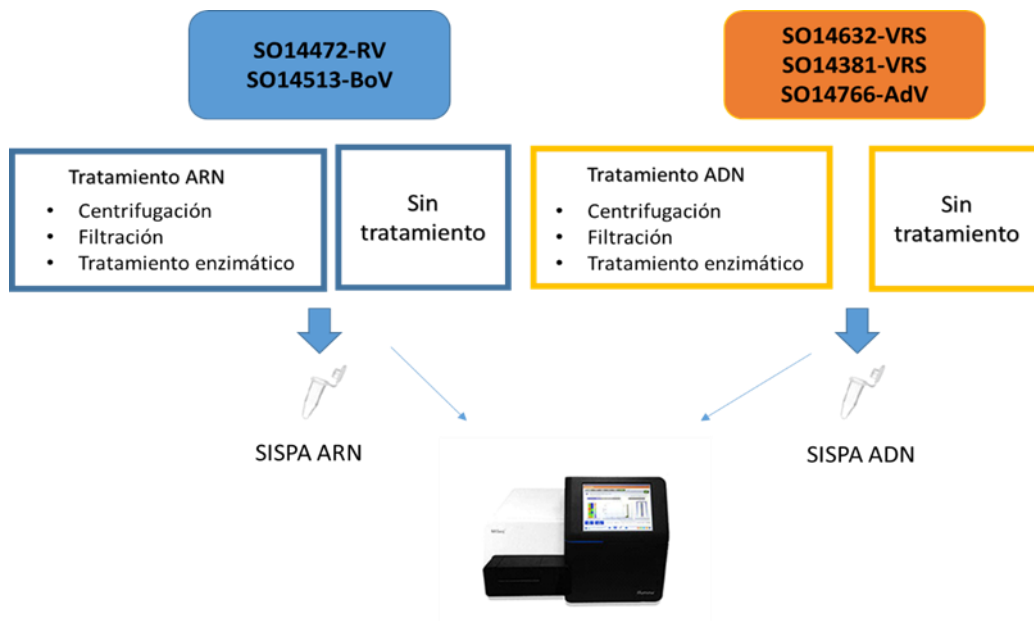
Durante el desarrollo del objetivo I fueron probados diferentes métodos específicos para el estudio de virus aplicados a las técnicas NGS.

El estudio del impacto del enriquecimiento viral en las muestras a analizar por NGS ha sido descrito previamente en trabajos realizados con diferentes tipos de muestras [179–186]. En la mayoría de estos trabajos la combinación de centrifugación a baja velocidad, filtración y digestión enzimática fue el método más empleado, por ello en este objetivo se desarrolló la adaptación de estos métodos a las características de la muestra respiratoria con el fin de comprobar su efecto en la secuenciación de viroma ADN y ARN del tracto respiratorio.

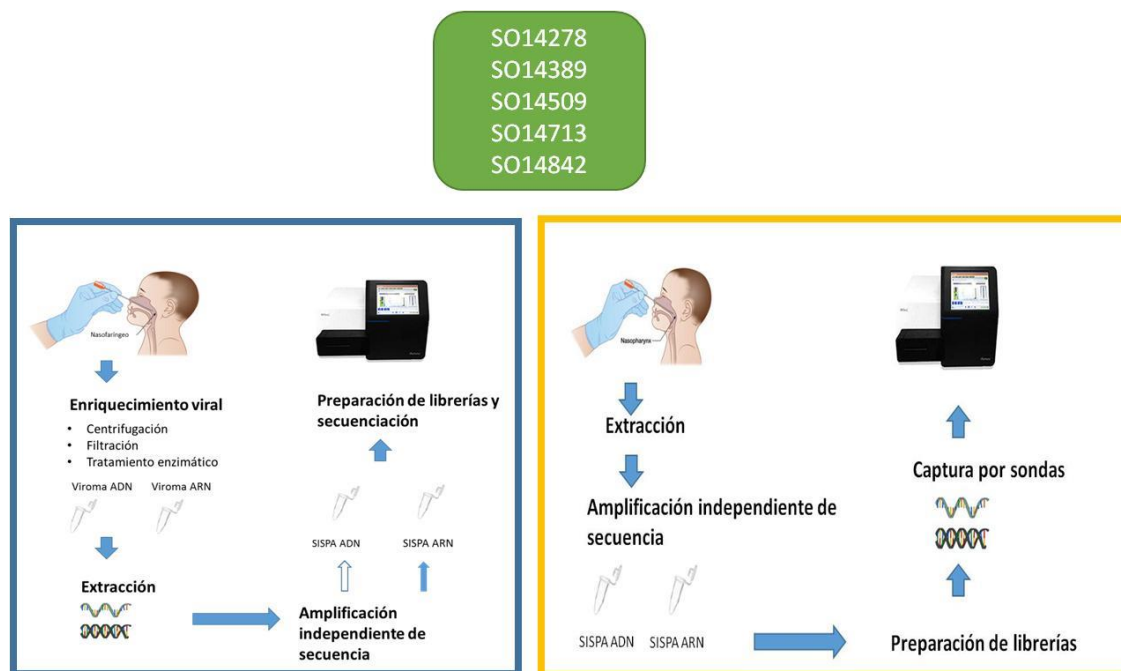
Las muestras respiratorias empleadas en el objetivo I procedían de pacientes pediátricos que formaban parte de los proyectos de investigación PI15CIII/00028 y PI18CIII/00009 (Materiales. Apartado 3.3.1). Estas muestras respiratorias fueron previamente analizadas con una batería de PCRs múltiples tiempo real de caracterización de virus respiratorios (Métodos. Apartado 4.12), la descripción de estas muestras y su diagnóstico está reflejado en la Tabla 3 (Página 62).

Las muestras respiratorias descritas fueron empleadas en el desarrollo de los 2 apartados que componen este objetivo: 1. la estandarización de métodos de enriquecimiento viral en muestras respiratorias y 2. la normalización del protocolo de amplificación independiente de secuencia con un cebador único para viroma ADN (SISPA ADN).

Los métodos empleados en la estandarización de métodos de enriquecimiento viral se basaron en dos aproximaciones diferentes: el flujo de trabajo descrito para viroma (Métodos. Apartado 4.5) con y sin enriquecimiento por 3 métodos como se representa en la Figura 11 y el método de enriquecimiento con el panel de captura por sondas *HyperCap Target Enrichment* (Métodos. Apartado 4.9) como se representa en la Figura 12.



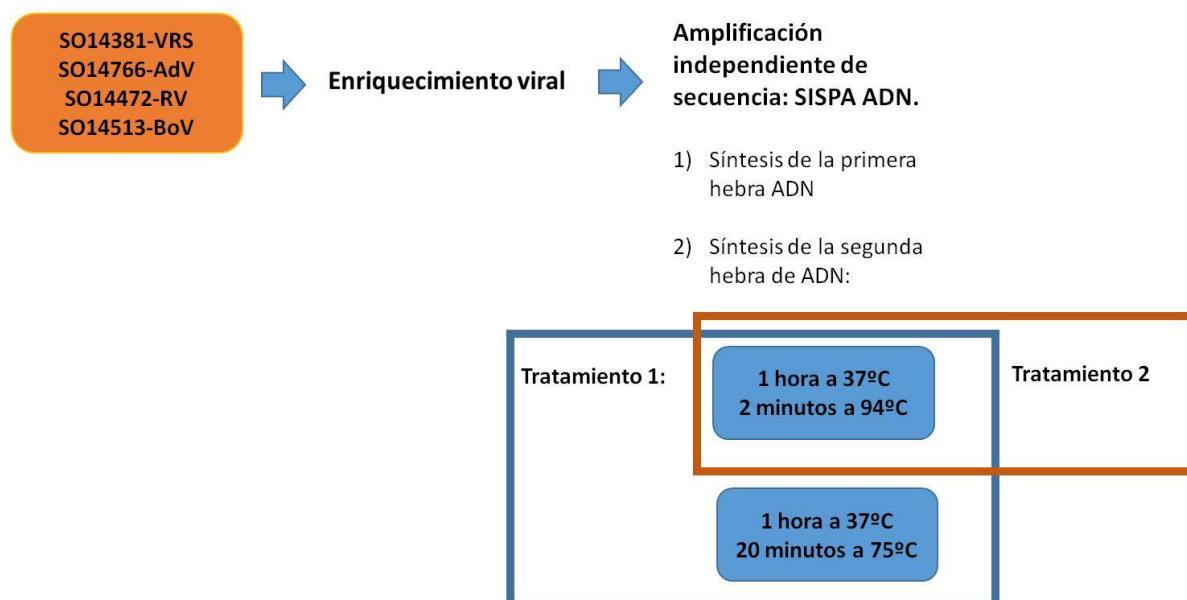
**Figura 11:** Flujo de trabajo sobre las muestras respiratorias pediátricas para comparar el número de fragmento de genomas virales obtenidos de las muestras enriquecidas y no enriquecidas.



**Figura 12:** Flujo de trabajo sobre las muestras pediátricas para comparar el número de fragmentos de genomas virales obtenidos en las muestras enriquecidas por los 3 métodos y las enriquecidas por captura de sondas

En el apartado de normalización del protocolo SISPA fueron evaluados dos puntos del protocolo:

1. el efecto de la aplicación de dos incubaciones consecutivas con el fragmento *Klenow* 3'→5' *exopolimerasa* en el protocolo de SISPA ADN y 2. la mezcla del producto de amplificación de SISPA ADN y ARN de una misma muestra respiratoria para su secuenciación conjunta.



**Figura 13:** Flujo de trabajo empleado en el estudio de la aplicación de la enzima Klenow exopolimerasa en una o dos incubaciones.

La preparación de librerías se realizó siguiendo el protocolo del kit *Nextera XT* (Métodos. Apartado 4.8), excepto cuando fue empleado el enriquecimiento de captura por sondas para el que se empleó el kit *Kapa Hyper Plus* de preparación de librerías (Materiales. Apartado 3.5.2).

Todas las muestras analizadas fueron secuenciadas usando el secuenciador *MiSeq* (Materiales. Apartado 3.5.4) y analizadas bioinformáticamente siguiendo el método descrito en el apartado 4.10 de Métodos en las páginas 88-89 empleando la plataforma *Virus TAP* [24] y las herramientas de análisis empleadas en USAMRIID.

Los resultados obtenidos en este objetivo reflejaron la importancia del enriquecimiento viral de las muestras antes de la secuenciación por NGS y la posibilidad de aplicar de manera exitosa dos métodos diferentes de enriquecimiento viral en la muestra respiratoria.



Además, fue demostrada que la mezcla de las reacciones de amplificación de las PCRs SISPA no conlleva a la pérdida de sensibilidad en la detección de los virus presentes en la muestra respiratoria reduciendo el número de librerías a realizar a la mitad.

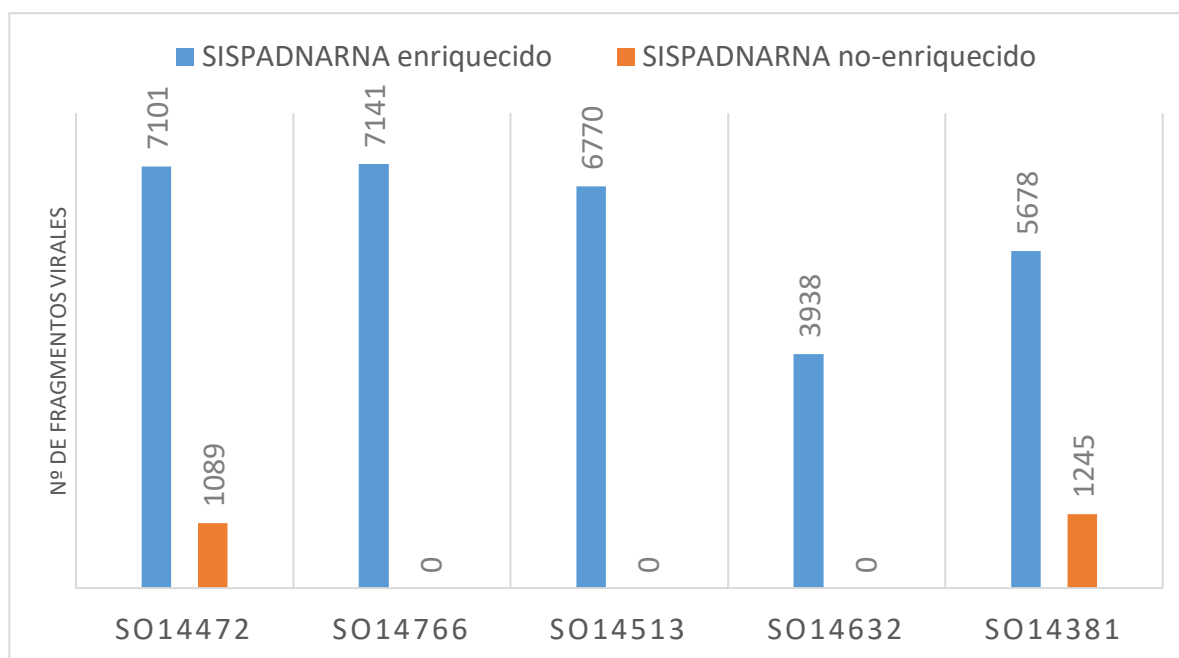
### 5.1.1. Estandarización de los métodos de enriquecimiento viral

Con el fin de comparar el número de fragmentos de genomas virales (o *contigs*) obtenidos tras el análisis de viroma ADN y ARN entre las muestras respiratorias enriquecidas por 3 métodos y sin enriquecer, realizamos el flujo de trabajo descrito en la Figura 11 sobre 5 muestras respiratorias (SO14472, SO14766, SO14513, SO14632 y SO14381) por duplicado, sin enriquecimiento viral y con enriquecimiento viral.

El análisis bioinformático de las secuencias obtenidas reveló que el número de fragmentos virales obtenidos tras el enriquecimiento por 3 métodos de las muestras respiratorias fue mayor que los fragmentos obtenidos en las muestras sin enriquecimiento viral (Figura 14. Página 117). De un total de 12,3 millones de lecturas obtenidas, el 25% de los fragmentos de genomas identificados pertenecían a secuencias virales tras el enriquecimiento viral de las muestras frente al 3% de fragmentos obtenidos en las muestras sin enriquecimiento viral. Tras realizar un análisis estadístico aplicando el test de la Chi-cuadrado observamos que la diferencia entre los resultados analizados es significativa obteniéndose un p-valor menor de 0,01.

En el caso de 3 de las 5 muestras analizadas no fue detectado ninguno de los virus presentes cuando las muestras no eran enriquecidas viralmente por 3 métodos. De estas 3 muestras negativas, 2 de ellas fueron previamente caracterizadas como positivas para virus ADN (SO14766 para AdV y SO14513 para BOV) y una de ellas para virus ARN (SO14632 para VRS), probando que la falta de enriquecimiento afecta a la secuenciación de ambos tipos viroma.

No solo el número de fragmentos resultó ser mayor tras la realización del enriquecimiento por 3 métodos sino que la secuencia obtenida de los virus detectados fue de mayor tamaño también. En la Tabla 15 se muestran los resultados obtenidos tras el análisis bioinformático comparando el tamaño de secuencia de los virus detectados en muestras que han sido enriquecidas y las que no. En el caso del tamaño de secuencia de los virus detectados en las muestras no enriquecidas fueron pequeños fragmentos de secuencia o *contigs* con poca profundidad de lectura.



**Figura 14:** Representación gráfica del número de fragmentos virales obtenidos tras el análisis de viroma ADN y ARN de las muestras enriquecidas y no enriquecidas

Identificación de muestra	Tamaño de secuencia del virus en muestra enriquecida	Tamaño de secuencia del virus en muestra no enriquecida
SO14472	RV (667 pb)	RV (95 pb)
SO14766	ADV (350 pb)	N/A
SO14513	BoV (954 pb)	N/A
SO14632	VRS (567 pb)	N/A
SO14381	VRS (437 pb)	VRS (125 pb)

**Tabla 15:** Identificación de la muestra analizada y tamaño de secuencia de los virus obtenidos en muestra respiratoria enriquecida y no enriquecida.

Otro de los métodos de enriquecimiento viral disponible es el método de captura por sondas que previamente se ha empleado en el enriquecimiento de virus respiratorios en otros trabajos publicados [151,152] y considerado un método más efectivo para el incremento de la proporción viral en la muestra [151]. Debido a la falta de recuperación de de secuencias completas de los virus analizados previamente empleando el enriquecimiento por los 3 métodos, en este apartado se realizó un ensayo para comparar los tamaños de secuencia de virus obtenidos empleando la captura por sondas y el enriquecimiento por 3 métodos.

Las muestras empleadas para este análisis fueron muestras positivas para RV previamente caracterizadas. Estas muestras se identificaron como SO14278, SO14839, SO14509, SO14713 y SO14842.

El flujo de trabajo aplicado en estas muestras es reflejado en la Figura 12. Por un lado el flujo de trabajo del estudio de viroma y la preparación de librerías empleando el kit *Nextera XT* (Materiales. 3.5.2), y por el otro, el método de enriquecimiento con el panel de captura por sondas *HyperCap Target Enrichment* (Materiales. Apartado 3.5.2), empleando el kit de preparación de librerías *Kapa Hyper Plus* (Figura 12. Página 114) (Materiales. 3.5.2.).

Identificación	Tamaño de secuencia tras enriquecimiento por 3 métodos	Tamaño de secuencia tras enriquecimiento empleando captura por sondas
SO14278	6393 pb	8268 pb
SO14389	N/A	N/A
SO14509	2100 pb	5768 pb
SO14713	3669 pb	7650 pb
SO14842	1630 pb	4771 pb

**Tabla 16:** Identificación de las muestras analizadas y tamaño de secuencia de los virus obtenidos en muestra enriquecida por 3 métodos y enriquecida mediante captura por sondas.

Los resultados obtenidos tras el análisis bioinformático (Tabla 16) mostraron que el tamaño de secuencia obtenido de los RV analizados es mayor cuando se empleó el método de enriquecimiento mediante captura por sondas con el panel de captura *HyperCap Target Enrichment* (Materiales. Apartado 3.5.2). Además, en el caso de las muestras SO14278 y SO14713, los RV detectados se obtuvo la secuencia completa de dichos RV empleando la captura por sondas.

La muestra respiratoria SO14839 pese a ser positiva en la PCR múltiple tiempo real Respiratorios 2 (Métodos. Apartado 4.12.1) no pudo ser secuenciada por NGS.

### 5.1.2. Normalización del protocolo de amplificación independiente de secuencia con un cebador único para viroma ADN

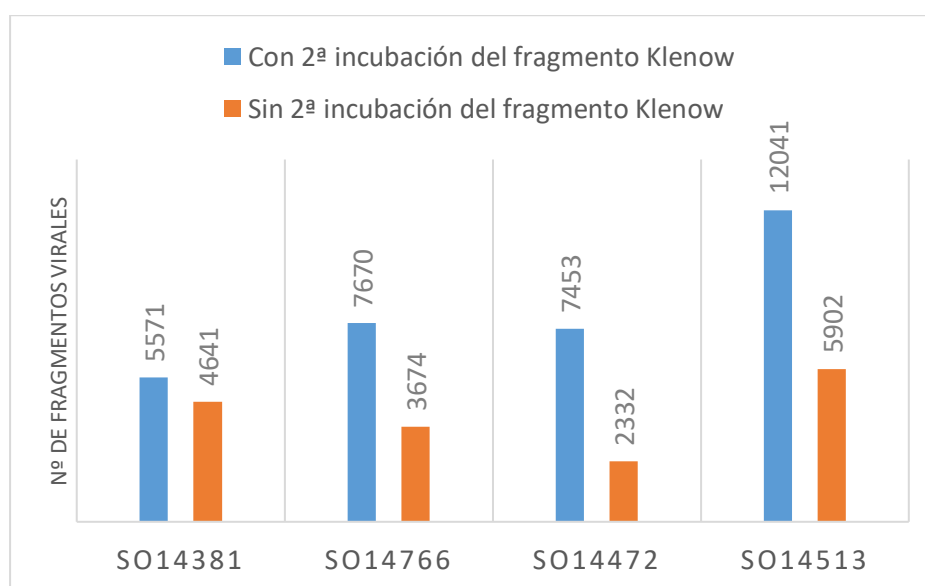
Como ha sido descrito en el flujo de trabajo del estudio de viroma tras el enriquecimiento viral se realizó la amplificación SISPA que permitía la amplificación de los diversos genomas virales presentes en la muestra de estudio.

En el protocolo de viroma ADN (Métodos. Apartado 3.5.2.1) descrito previamente fue reflejada la existencia de dos incubaciones consecutivas del extracto de la muestra enriquecida con el fragmento Klenow 3'→5' exopolimerasa (Materiales. Apartado 3.5.1) para la formación de la doble cadena de ADN. Con el fin de conocer la posibilidad de optimizar el protocolo, fue comprobado si únicamente con una única incubación del fragmento Klenow 3'→5' exopolimerasa con el extracto se producía la síntesis exitosa de la doble cadena.

Las muestras empleadas y el flujo de trabajo fueron descritos en la Figura 13.

El análisis realizado fue la comparación entre el número de fragmentos virales, tras realizar el protocolo de trabajo con 2 incubaciones de la muestra con el fragmento Klenow 3'→5' exopolimerasa o simplemente con 1 incubación.

Los resultados obtenidos reflejados en la Figura 15 mostraron la importancia de la realización de la segunda incubación con el fragmento Klenow 3'→5' exopolimerasa. En 3 de las 4 muestras analizadas el número de fragmentos de virus obtenidos era muy superior tras la aplicación de la doble incubación durante la realización del SISPA ADN. En el caso de todas las muestras respiratorias analizadas, a excepción de la muestra SO14381, el número de fragmentos virales de las muestras con segunda incubación de la enzima Klenow exopolimerasa casi triplicaba a las que solo tuvieron una única incubación. Tras realizar un análisis estadístico aplicando el test de la Chi-cuadrado observamos que la diferencia entre los resultados analizados es significativa obteniéndose un p-valor menor de 0,01.



**Figura 15:** Número de fragmentos virales obtenidos con la 2ª incubación de la enzima *Klenow exopolimerasa* y sin la 2ª incubación.

Uno de los problemas generados por la diferencia entre los protocolos de SISPA para ADN y ARN es la producción de dos mezclas de reacción diferentes que conlleva a la preparación del doble de librerías.

Debido a esto en este apartado se realizó la comparación de la representación viral obtenida tras la secuenciación NGS de las librerías procedentes de los productos de SISPA ADN y ARN por separado y mezclados equimolarmente.

Tras la medida de la concentración de ADN en los productos de amplificación de ambos métodos SISPA se ajustó cada producto de amplificación a 500 ng de concentración y se mezclaron.

Los resultados obtenidos tras el análisis bioinformático (Tabla 17) mostraron que los virus detectados en las librerías procedentes del producto de los SISPA de forma individual se veían también reflejados en los resultados obtenidos en las librerías procedentes de la mezcla.

Tras la confirmación de que no se perdía detección de virus en la mezcla de los productos de amplificación de SISPA, todos los procedimientos posteriores se realizaron empleando la mezcla para la preparación de librerías.

Identificación	Virus detectados SISPA ADN	Virus detectados SISPA ARN	Virus detectados mezcla SISPA ADN/ARN
SO14278	N/A	RV	RV
SO14389	TTV	N/A	TTV
SO14509	TTV	RV	TTV y RV
SO14713	TTV	RV	TTV y RV
SO14842	N/A	RV	RV

**Tabla 17:** Resultados de la prueba de la comparación entre los virus detectados en viroma ADN y viroma ARN y su mezcla equimolar

## DISCUSIÓN

Este primer objetivo se centra en la puesta a punto necesaria para comenzar a desarrollar la caracterización de viroma con la garantía de estar empleando el método más eficaz y que mejor se adapta a las características de la muestra respiratoria.

En el desarrollo de estrategias universales de enriquecimiento debe ser tenido en cuenta que los virus que conforman el viroma respiratorio presentan múltiples tamaños y características estructurales específicas, por lo que en el empleo de algunos de los métodos de enriquecimiento

viral se puede perder información acerca de los virus presentes o crear sesgos que sobreexpresen secuencias de algunos de los virus detectados frente a las secuencias de otros virus [187,188].

La elección de la técnica de enriquecimiento viral a utilizar se basa en los métodos referenciados en trabajos previos de metagenómica [9,180,189–191]. Hay que tener en cuenta que estos trabajos previos son empleados en muestras de diferente procedencia biológica y de que no todas las técnicas descritas tienen el mismo efecto sobre los diferentes virus presentes, incluso en el mismo tipo de muestra [9,192]. Por esta razón, la estandarización de los métodos empleados en las muestras respiratorias disponibles es imprescindible para un buen rendimiento de los métodos NGS en el estudio de viroma respiratorio.

El primer apartado de este objetivo pretende conocer el impacto del enriquecimiento y del tipo de enriquecimiento en la caracterización del viroma respiratorio y la obtención de genoma completo de los virus respiratorios. Los resultados obtenidos mostraron que el enriquecimiento es crítico para la detección de virus en las muestras respiratorias. La falta de enriquecimiento conlleva falta de sensibilidad de la secuenciación, no detectándose virus presentes en la muestra y en el caso de ser detectados presentando un tamaño de secuencia mucho menor que los de las secuencias obtenidas tras el tratamiento de enriquecimiento (Figura 14 y Tabla 15. Página 117).

El empleo de filtrados con diferente tamaño de poro en los protocolos de enriquecimiento es uno de los puntos críticos para la detección de virus de gran tamaño. Uno de los tamaños de poro empleados en trabajos previos es el poro de 0,22  $\mu\text{m}$ . Se ha descrito que el empleo de este tamaño de poro dificulta una buena representación de los virus de gran tamaño como los virus herpes [191,193–195]. En el caso de esta tesis, se realizó un doble filtrado con poros de 5  $\mu\text{m}$  y 0,45  $\mu\text{m}$  (Métodos. Apartado 4.5.1). El empleo de este doble filtrado permitió la secuenciación de un virus ADN de gran tamaño como es el AdV detectado en la muestra SO14766 (Tabla 16. Página 118).

Actualmente están siendo empleando métodos de captura por sondas que han demostrado ser efectivos a la hora de incrementar el número de lecturas virales y el tamaño de secuencia obtenido de los virus detectados [151]. Las ventajas de la aplicación de este método son la eliminación de filtros físicos que evita sesgos por morfología y tamaño, además de que la sensibilidad de este método es comparable a la obtenida mediante las PCRs en tiempo real específicas por virus [151].

El empleo de muestras respiratorias pediátricas con RV, detectado previamente por PCR tiempo real Respiratorios 2 [159], permitió la evaluación de la aplicación del enriquecimiento mediante captura por sondas en muestras respiratorias pediátricas.

Los resultados mostraron que la detección viral fue igual empleando la captura por sondas que el enriquecimiento por 3 métodos (Tabla 16. Página 118). Sin embargo, empleando el método de enriquecimiento por captura por sondas los tamaños de secuencia obtenidos de los virus detectados fueron mayores, pudiendo obtenerse la secuencia genómica completa como en el caso del RV detectado en la muestra SO14278. A pesar de mejorar el tamaño de secuencia obtenido de los diferentes virus, el empleo de enriquecimiento mediante captura por sondas, con el panel de captura HyperCap Target Enrichment (Materiales. Apartado 3.5.2) asociado a la realización de librerías empleando el kit Kapa Hyper Plus (Materiales. Apartado 3.5.2), presenta dos inconvenientes principales: 1. la dificultad y duración de su protocolo y 2. la posibilidad de pérdida de detección de virus divergentes a las sondas diseñadas o incluso nuevos virus.

De acuerdo a las características de los protocolos empleados (Métodos. Apartado 4.9), la combinación de este método de elaboración de librerías y del enriquecimiento mediante captura por sondas es un proceso más laborioso y largo en comparación con los tiempos de trabajo estimados empleando el enriquecimiento por 3 métodos (centrifugación, filtración y digestión enzimática) y la preparación de librerías con el kit Nextera XT (Materiales. Apartado 3.5.2).

Además, en el caso de los protocolos para el estudio de viroma y el descubrimiento de patógenos, el enriquecimiento mediante captura por sondas puede dificultar el hallazgo de nuevos virus o virus con baja homología frente a los virus empleados en el diseño de las propias sondas empleadas en la captura del ADN problema. En el caso de virus con alta tasas de mutación, este método puede perder sensibilidad al no reconocer secuencias virales con una homología menor del 75% con las sondas diseñadas [151]. Por ello, aunque ambos métodos tienen sus ventajas e inconvenientes, son útiles en el estudio de viroma y se deben elegir en función del objetivo deseado. Cuando el objetivo sea la detección de los virus presentes o el descubrimiento de nuevos virus en la muestra respiratoria, sin necesidad del estudio de secuencia completa, el empleo del enriquecimiento por 3 métodos propuesto en esta tesis es un método útil y eficaz. En el caso de que el objetivo sea la obtención de la secuencia completa de los virus presentes en la muestra respiratoria, el método óptimo a emplear sería el enriquecimiento mediante captura por sondas.

Independientemente del método de enriquecimiento viral escogido, el empleo de una técnica de amplificación aleatoria puede ayudar a aumentar el material genético presente en la muestra, por lo que si tras el enriquecimiento hay una mayor proporción viral usando este método se incrementará el número de fragmentos virales obtenidos por NGS en el estudio de viroma [196].

Con el fin de reducir el tiempo de trabajo empleado en la realización del protocolo mediante SISPA ADN se comprobó la reducción de las reacciones de incubación del extracto de la muestra enriquecida con la enzima Klenow exopolimerasa. Estas incubaciones presentaban las mismas condiciones de temperatura y tiempo. Finalmente, la realización de las dos incubaciones propuestas en el protocolo mostró que el número de fragmentos virales obtenidos fue mayor que el número de fragmentos virales obtenidos tras la realización de una única incubación (Figura 15. Página 119).

Otra de las optimizaciones propuestas en el desarrollo de los métodos SISPA ADN y ARN fue la eliminación del duplicado de la misma muestra en la preparación de librerías, librería del producto de SISPA ADN y del producto de SISPA ARN, para el incremento del número de muestras respiratorias de pacientes a analizar por carrera NGS, sin perder profundidad de lectura. Trabajos previos mostraron que la realización de estas mezclas de producto de reacción mejoraban uno de los sesgos producidos por el empleo de este método, debido a la distribución desigual de las zonas de unión de los cebadores y que puede producirse una sobresecuenciación de algunos fragmentos de genoma, afectando a la detección de los virus presentes en baja proporción en la muestras [187,188]. El empleo de la mezcla de productos de reacción puede favorecer un patrón de cobertura más uniforme [197].

Los resultados obtenidos tras el análisis de las secuencias obtenidas por NGS de las mezclas de las dos reacciones de amplificación SISPA ADN y ARN, no supuso una pérdida de sensibilidad pudiendo ser detectados los mismos virus tras la secuenciación NGS de las muestras de reacción por separado y mezcladas (Tabla 17. Página 120).

Como conclusión, el empleo de métodos de enriquecimiento viral en las muestras, previos a la secuenciación por NGS, y la comparación de dos métodos diferentes de enriquecimiento viral de la muestra respiratoria, desarrollados en esta tesis, permiten su aplicación en este tipo de muestras y la consecución del resto de los objetivos planteados en esta tesis.



## **5.2 OBJETIVO II: Aplicación de los métodos desarrollados en el estudio de la composición del viroma respiratorio y las relaciones ecológicas de sus integrantes**

Este objetivo se ha centrado en dos líneas diferentes de trabajo que, por el interés que han presentado en el estudio de la infección respiratoria, se van a describir de manera independiente:

1. estudio del viroma respiratorio de pacientes pediátricos: presencia de RV y su posible relación con otros virus y 2. descripción por primera vez, de un nuevo virus ADN circular detectado en el tracto respiratorio.

### **5.2.1 Línea de trabajo 1: Estudio del viroma respiratorio de pacientes pediátricos. Presencia de RV y su posible relación con otros virus**

En esta línea de trabajo se desarrolló el estudio de la comunidad viral del tracto respiratorio de pacientes pediátricos con el fin de obtener más información sobre las características de la infección respiratoria asociada a virus además de conocer las implicaciones de los virus detectados en los pacientes infectados.

Las características de las muestras empleadas fueron descritas en el Apartado 3.3.2 de Materiales. El tipo de muestra empleada fue ANF de pacientes pediátricos procedentes de los proyectos de investigación PI15CIII/00028 y PI18CIII/00009 (Figura 2, Tabla 4. Páginas 63-67), estas muestras fueron previamente analizadas empleando las PCRs múltiples tiempo real BRQ, Respiratorios 2 y Gripe A, B y C (Métodos. Apartado 4.12).

En esta línea de trabajo se abordaron los siguientes puntos: 1. caracterización del viroma de pacientes con infección respiratoria pero sin detección de virus por PCR; 2. caracterización de viroma de individuos control sanos; 3. el estudio de la presencia de RV en pacientes con IRA y 4. relación de RV con otros virus presentes en el tracto respiratorio.

La caracterización de viroma se realizó mediante el análisis de dos tipos de muestras: pacientes sanos y pacientes con IRA que fueron negativos en las PCRs de caracterización virológica (Figura 16). Los individuos sanos, que conformaban el grupo control, fueron seleccionados para ser comparables con los pacientes con IRA y con resultados negativos. El grupo de control presentaba

una edad similar, provenían de la misma área geográfica y sus muestras clínicas también se tomaron sistemáticamente durante la misma ventana de tiempo.

En ambos estudios fue empleado el protocolo de trabajo para el estudio de viroma (Métodos. Apartado 4.5) y la preparación de librerías se realizó empleando en kit *Nextera XT* (Materiales. Apartado 3.5.2) para su secuenciación empleando el secuenciador *MiSeq* (Materiales. Apartado 3.5.4).

Para el estudio de la presencia de RV en pacientes con IRA y su relación con otros virus presentes en el tracto respiratorio fueron empleadas muestras de seguimiento de pacientes con IRA a los que se les detectó RV (Materiales. Apartado 3.3.2). A las muestras de análisis se les realizó el método de enriquecimiento con el panel de captura por sondas *HyperCap Target Enrichment* (Materiales. Apartado 3.5.2) y la preparación de librerías con el kit *Kapa Hyper Plus* (Métodos. Apartado 4.9) para su secuenciación con el secuenciador *MiSeq* (Materiales. Apartado 3.5.4). Para el estudio de virus circulares en las muestras de seguimiento fue empleado el método RCA (Métodos. Apartado 4.6) (Figura 16).

El análisis bioinformático en todas las muestras empleadas en todos los objetivos fue realizado con la plataforma Virus TAP y el análisis de datos de realizado con el método empleado en USAMRIID descrito en el apartado 4.10.1 de Métodos.

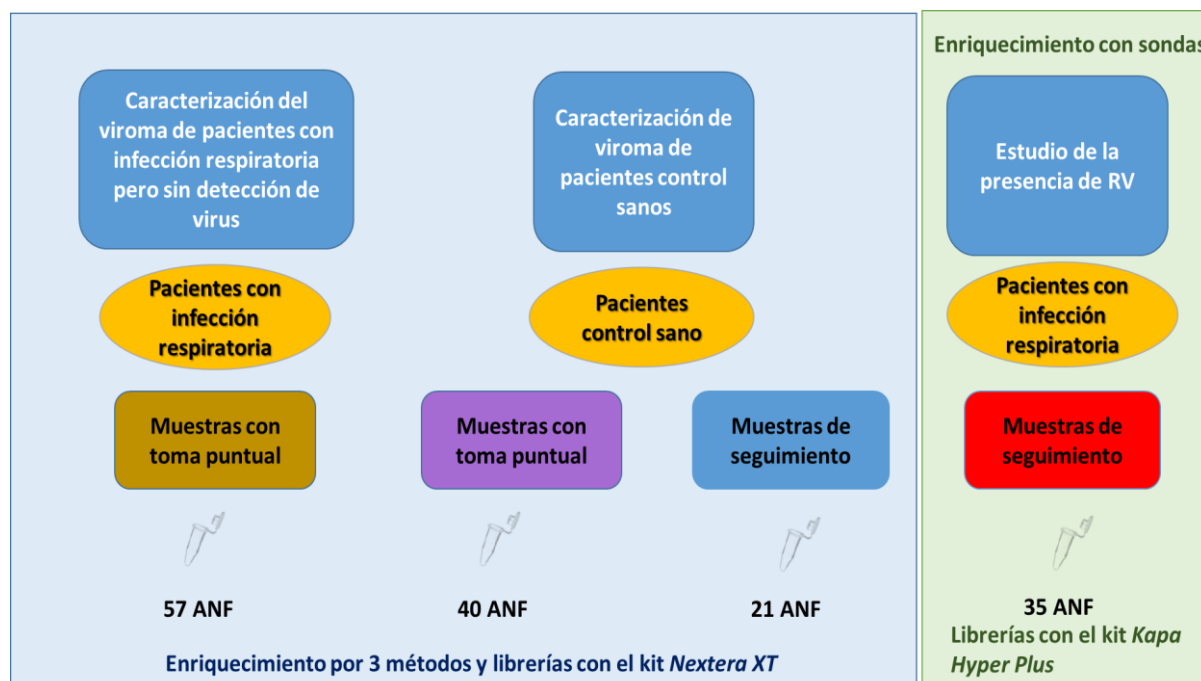
Los resultados obtenidos en este objetivo permitieron la caracterización del viroma de pacientes pediátricos sanos y con infección respiratoria que no fueron diagnosticados empleando métodos moleculares de diagnóstico.

El análisis de las muestras de 61 pacientes sanos mostró la ausencia de patógenos respiratorios conocidos, confirmando que en ausencia de infección no hay presencia de virus patógenos en los ANFs analizados.

De los 57 pacientes con síntomas clínicos de infección respiratoria sin diagnóstico fueron caracterizados los virus patógenos causantes de la infección en 35 de ellos (61%) empleando el flujo de trabajo de viroma respiratorio. En 22 de estas muestras (39%) no se halló ningún virus patógeno conocido.

El estudio de RV permitió caracterizar a los virus presentes a lo largo del tiempo en muestras respiratorias de seguimiento en pacientes con infecciones recurrentes de estos virus. La

presencia de TTV fue más alta en pacientes con infección respiratoria frente a pacientes sanos y se detectó fundamentalmente en pacientes hospitalizados y con infecciones recurrentes.

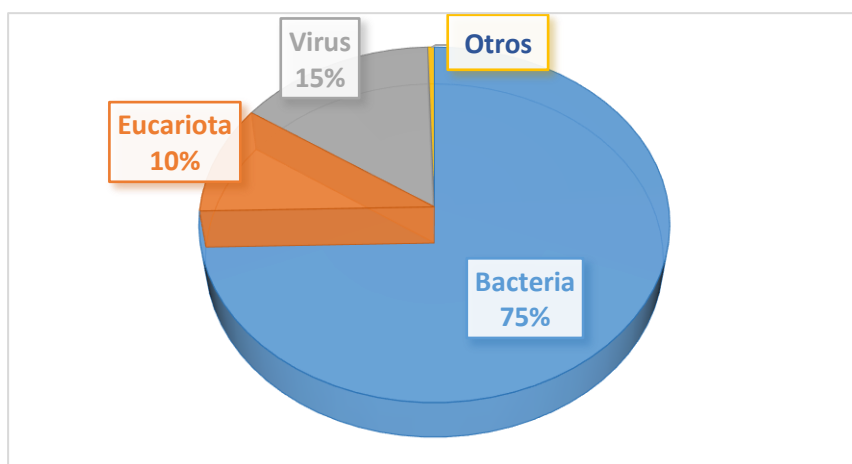


**Figura 16:** Método de enriquecimiento, de preparación de librerías y muestras empleadas en caracterizar el viroma de pacientes sanos (40+21 ANF), pacientes con IRA (57 ANF) y en el estudio de la presencia de RV en muestras de seguimiento (35 ANF).

#### 5.2.1.1. Virus identificados en pacientes con infección respiratoria empleando el flujo de trabajo para el estudio de viroma por NGS

En este apartado fueron analizados 57 ANFs procedentes de pacientes pediátricos cuyo historial clínico refería síntomas producidos por infección respiratoria (Tabla 4. Páginas 63-67). Estos ANFs fueron previamente analizados con las PCRs de caracterización de virus respiratorios realizadas en el laboratorio (Métodos. Apartado 4.12) resultando negativos para todos los virus respiratorios estudiados.

El análisis bioinformático (Métodos. Apartado 4.10) realizado tras la secuenciación NGS de los 57 ANF mostró que se habían producido un total de 24,6 millones de lecturas procedentes de las 57 muestras. Los fragmentos de secuencias genómicas taxonómicamente identificados empleando BLASTn (Materiales. Apartado 3.5.5) fueron 35,666 en total, de los cuales correspondían 5117 fragmentos a secuencias de virus, representando el 15% del total de fragmentos identificados por BLASTn (Figura 17 en gris).

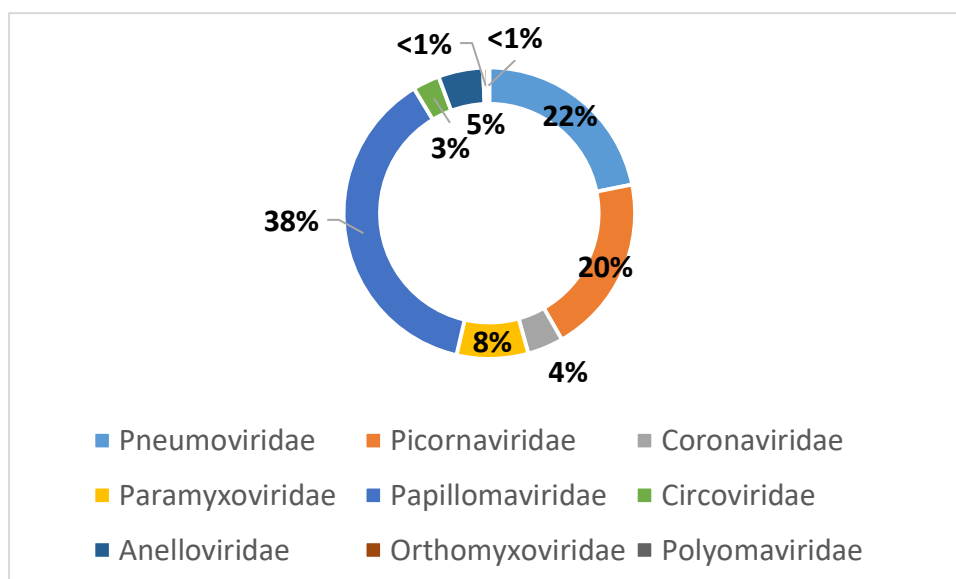


**Figura 17:** Abundancia relativa de los fragmentos obtenidos tras el análisis de las muestras respiratorias de los pacientes pediátricos con IRA pero sin identificación de patógeno viral.

Dentro de estos 5117 fragmentos correspondientes a virus, el 13% (667 fragmentos con 2,6 millones de lecturas) correspondieron a virus que infectan a organismos eucariotas, mientras el 87% (4450 fragmentos con 4.1 millón de lecturas) correspondieron a bacteriófagos.

Entre los fragmentos asignados a virus que infectan organismos eucariotas se identificaron 9 familias virales: *Papillomaviridae*, *Pneumoviridae*, *Picornaviridae*, *Paramyxoviridae*, *Anelloviridae*, *Coronaviridae*, *Circoviridae*, *Orthomyxoviridae* y *Polyomaviridae* (Figura 18).

Teniendo en cuenta que el análisis se centraba en la búsqueda de patógenos virales que explicaran el origen de la infección respiratoria, se detectaron 340 fragmentos de genomas que correspondían a virus patógenos conocidos (Tabla 18). La mayoría de estas secuencias correspondieron a los virus VRS y RV. La presencia de estos virus fue confirmada por las PCRs múltiples tiempo real BRQ y Respiratorios 2 (Métodos. Apartado 4.12.1).



**Figura 18:** Identificación taxonómica de los fragmentos obtenidos tras el análisis de las muestras respiratorias de los pacientes pediátricos con IRA pero sin identificación de patógeno viral.

Resultados del análisis de las muestras con IRA			
Virus	Fragmentos	Lecturas	Muestras positivas
VRS	141	1,683,686	15
MNV	3	471,167	1
EV	18	206,104	4
CoV	29	88,760	3
RV	90	64,659	12
RuV	56	29,595	4
Gripe B	3	18	2

**Tabla 18:** Patógenos respiratorios virales detectados tras el análisis de las muestras respiratorias con IRA pero sin identificación de patógeno viral. RuV: Rubulavirus.

Gracias al empleo de este método pudieron ser identificados los patógenos virales causantes de la infección en 35 de los 57 pacientes pediátricos estudiados (61%). Entre los virus detectados en estas muestras se detectó el virus HRCiV, descrito posteriormente y caracterizado por primera vez en esta tesis.

En 22 de las 57 muestras (39%) no se detectó ningún patógeno respiratorio viral conocido.

Muestra	Fragmentos / Lecturas			
	TTV	HRCiV	Papilomavirus	Poliomavirus 4
1	1 / 146	3 / 121	18 / 569	-
2	-	2 / 3	11 / 149	-
3	-	-	-	-
4	-	-	-	-
5	-	-	1 / 1	1 / 3
6	-	-	9 / 30	-
7	-	-	-	-
8	-	-	-	-
9	-	-	1 / 2	-
10	-	3 / 12	3 / 195	-
11	-	2 / 6	9 / 104	-
12	-	-	13 / 118	-
13	2 / 4	-	5 / 36	-
14	-	1 / 2	2 / 7	-
15	-	-	-	-
16	-	-	3 / 156	-
17	-	1 / 3	7 / 30	-
18	-	-	1 / 4	-
19	-	-	4 / 38	-
20	-	-	4 / 215	-
21	-	-	3 / 10	-
22	-	1 / 101	3 / 1,961	-

**Tabla 19:** Número total de fragmentos y lecturas de los virus detectados en las muestras de pacientes con IRA pero sin identificación significativa de patógeno viral.

Con el fin de asegurar que durante el análisis bioinformático no se estaba perdiendo información se realizó un análisis más profundo de los fragmentos no identificados empleando la herramienta *BLASTx* (Materiales. Apartado 3.5.5) para la búsqueda de homologías con proteínas virales conocidas. Tras este análisis no se obtuvieron resultados positivos significativos. Los fragmentos analizados presentaron identificaciones virales no significativas y relacionadas con otras

proteínas de bacterias con los mismos valores de significación, por lo que no se pudieron considerar coincidencias específicas a virus significativas.

Un segundo análisis se realizó sobre las secuencias que no formaron fragmentos, estas lecturas fueron estudiadas mediante alineamientos a pares con las secuencias depositadas en la base de datos *GenBank* (Materiales. Apartado 3.5.6) empleando el programa DIAMOND [174]. Tras este segundo análisis se detectaron bacteriófagos, TTV, el nuevo CRESS-DNA virus HRCiV y poliomavirus humano tipo 4 (Tabla 19. Página 129).

#### **5.2.1.2. Caracterización de viroma del grupo control sanos**

En este apartado fueron analizadas 61 muestras procedentes del grupo de los niños control sanos descritos en la Tabla 4 (Páginas 63-67), empleando el flujo de trabajo del estudio de viroma (Métodos. Apartado 4.5), con el fin de conocer y caracterizar los virus presentes y comparar los resultados con los obtenidos en el apartado anterior.

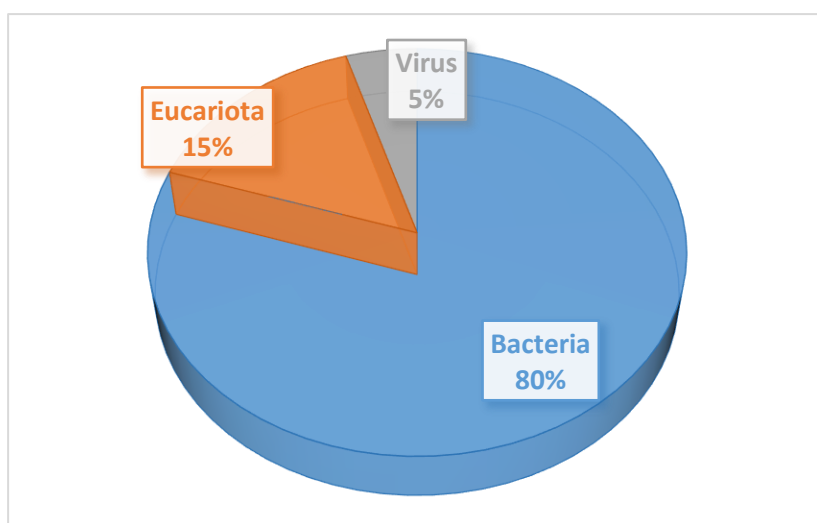
Estos individuos control sanos no presentaban síntomas de infección respiratoria ni les fue detectado ningún virus respiratorio en el análisis de sus muestras por las PCRs múltiples tiempo real de diagnóstico de virus respiratorios (Métodos. Apartado 5.12).

El análisis bioinformático (Métodos. Apartado 4.10) realizado tras la secuenciación NGS de las 61 muestras mostró la obtención de 40,1 millones de lecturas. Se obtuvieron 42.013 fragmentos de secuencias genómicas identificadas como virales empleando BLASTn (Materiales. Apartado 3.5.5) (Figura 19).

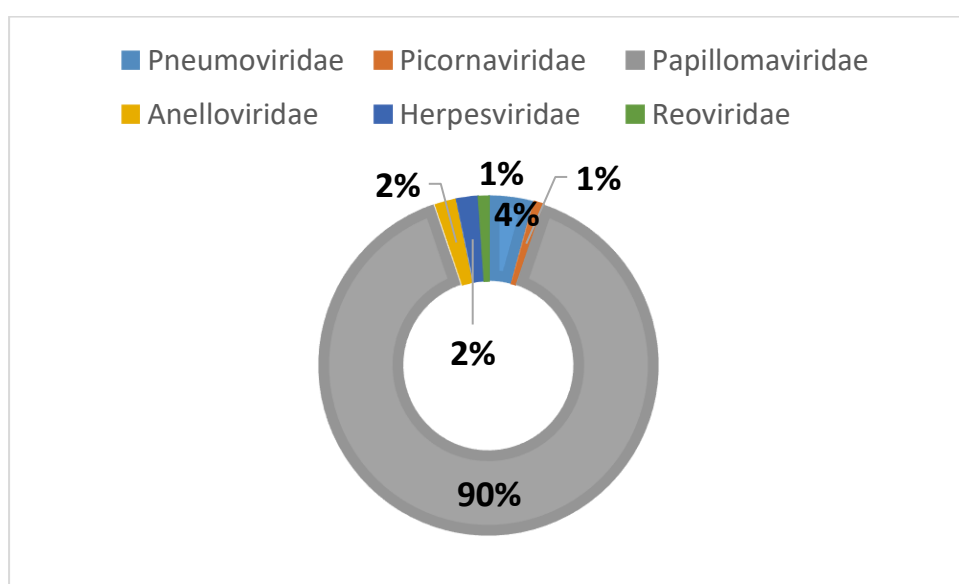
Dentro de estos 42.013 fragmentos, el 5% (1974 fragmentos) de los fragmentos de virus correspondían a virus que infectan a eucariotas y 95,3% a bacteriófagos.

Entre los fragmentos asignados a los virus que infectan eucariotas se encontraron virus de las familias *Papillomaviridae* (90%), *Pneumoviridae* (4%), *Anelloviridae* (2%), *Picornaviridae* (1%), *Herpesviridae* (2%) y *Reoviridae* (1%) (Figura 20).

Los fragmentos de virus de la familia *Papillomaviridae* fueron los más abundantemente secuenciados representando el 0,9% del total de fragmentos secuenciados en estas muestras. A pesar de estos datos, solo fue detectado un tipo de papilomavirus, el papilomavirus humano “Dysk 6” (GenBank: KX781285), presente únicamente en 11 de los 61 pacientes analizados.



**Figura 19:** Abundancia relativa de los fragmentos obtenidos tras el análisis de las muestras respiratorias de los individuos pediátricos control sanos



**Figura 20:** Identificación taxonómica de los fragmentos obtenidos tras el análisis de las muestras respiratorias de los individuos pediátricos control sanos

En el caso de estas muestras control solo fueron detectados dos virus patógenos respiratorios en 2 de los 61 ANFs: RV y HMNV (Tabla 20) confirmados por PCRs múltiples tiempo real BRQ y Respiratorios 2 (Métodos. Apartado 4.12.1)



Resultados del análisis de las muestras de pacientes control			
Virus	Fragmentos	Lecturas	Muestras positivas
HMNV	4	943,278	1
RV	2	5	1

**Tabla 20:** Número total de fragmentos y lecturas de los virus detectados en las muestras del grupo control sanos.

Como conclusión, en 59 de los ANFs analizados (97%) no se detectó presencia de virus patógenos respiratorios conocidos.

### 5.2.1.3 Caracterización de la presencia de RV en las muestras pediátricas

La infección por RV es una de las causas más comunes de hospitalización por IRA [198] y un factor de riesgo significativo para presentar sibilancias permanentes en pacientes pediátricos [51].

Gracias a la disponibilidad de muestras de seguimiento (Materiales. Apartado 3.2.2.) tanto de pacientes con IRA como de pacientes sanos fue posible el estudio de caracterización de RV en pacientes pediátricos a lo largo del tiempo.

Para este estudio fueron empleados 35 ANFs procedentes de 11 pacientes con infección respiratoria a los que previamente se les había detectado RV en todas las muestras empleadas mediante la PCR múltiple tiempo real Respiratorios 2 (Métodos. Apartado 4.12.1).

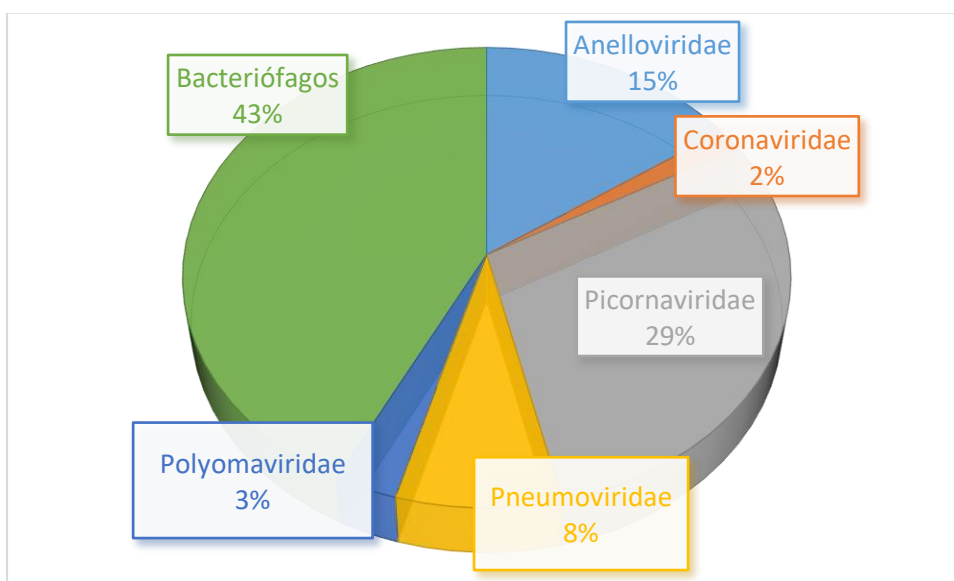
Tras el análisis bioinformático de las secuencias obtenidas por NGS de las muestras estudiadas, se pudieron secuenciar 27,2 millones de lecturas. Los fragmentos de secuencias genómicas taxonómicamente identificados como virus, empleando BLASTn (Materiales. Apartado 3.5.5), fueron 3,2 millones representando el 14% del total.

De los 3,2 millones de fragmentos virales obtenidos el 40% correspondían a bacteriófagos y el 60% restante (1,9 millones de fragmentos) a virus que infectan a eucariotas (Figura 21).

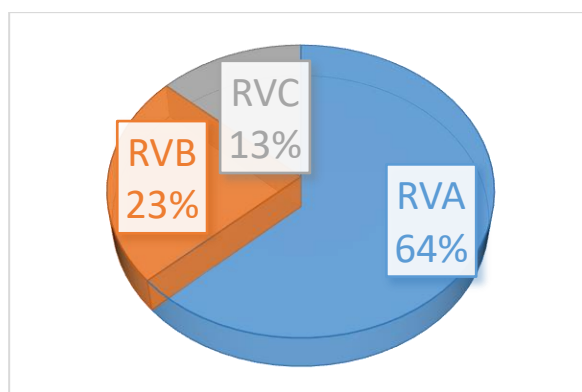
Se detectaron 5 familias de virus que infectaban a eucariotas: *Anelloviridae* (15%), *Coronaviridae* (2%), *Picornaviridae* (29%), *Pneumoviridae* (8%) y *Polyomaviridae* (3%).

La familia de virus con mayor número de fragmentos detectados fue la familia Picornaviridae, familia a la que pertenecen los RV.

Mediante el empleo de la PCR múltiple tiempo real Respiratorios 2 (Métodos. Apartado 4.12.1) no es posible determinar el tipo de RV detectado, sin embargo el empleo del enriquecimiento mediante captura por sondas y la posterior secuenciación NGS realizada en este apartado, permitió la detección de los RV y su tipado específico pudiendo ser determinado el porcentaje de cada grupo taxonómico descrito de RV en cada uno de los 35 ANF. Los análisis reflejaron una mayor presencia de RVA correspondiendo al 64% de los virus analizados, seguido de RVB con un 23% y finalmente RVC con un 13% (Figura 22).



**Figura 21:** Identificación taxonómica de los virus detectados tras el análisis de las muestras respiratorias de seguimiento de los pacientes pediátricos con RV.



**Figura 22:** Distribución del tipo de RV detectados en las muestras de seguimiento de pacientes pediátricos

El empleo de muestras de seguimiento permitió la caracterización de los RV detectados en los pacientes a lo largo del tiempo. De los 11 pacientes estudiados se obtuvieron fragmentos de RV

en 10 de ellos, abarcando periodos de tiempo de hasta 8 meses de estudio con presencia de RV en las muestras de seguimiento (Tabla 21). En el caso de los pacientes 127 y 128 fueron tomadas dos muestras en el mismo mes, con 10 días de diferencia en el caso del paciente 127 y con 8 días de diferencia en el paciente 128.

No fue posible realizar el seguimiento en todos los pacientes por falta de resultados en la obtención de secuencia suficiente que permitiera comparar los porcentajes de homología entre los virus detectados en las muestras de seguimiento del mismo paciente (paciente 124), o debido a que no fue posible comparar las secuencias de RV dado que los tipos de RV detectados fueron diferentes en cada una de las muestras demostrando diferentes episodios no relacionados (paciente 129).

Tiempo de toma de la muestra en meses								
Pacientes	1	2	3	4	5	6	7	8
121	RVA	N/A	RVA	N/A	RVA	N/A	N/A	N/A
122	RVA	RVA	N/A	N/A	N/A	N/A	N/A	N/A
124	RVA	RVA	N/A	N/A	RVB	N/A	N/A	N/A
125	RVB	N/A	RVB	N/A	N/A	RVA	N/A	N/A
126	RVA	RVA	RVA	RVA	N/A	N/A	N/A	N/A
127	RVA/RVA	RVB	N/A	N/A	N/A	N/A	N/A	N/A
128	RVA/RVA	RVC	N/A	RVC	N/A	N/A	N/A	N/A
129	RVB	N/A	RVC	N/A	N/A	N/A	RVB	N/A
130	RVA	N/A	N/A	N/A	RVB	N/A	N/A	N/A
131	RVA	N/A	N/A	N/A	N/A	N/A	N/A	RVA

**Tabla 21:** Línea temporal de la toma de muestras positivas para RV y los tipos de RV detectados en ellas

La caracterización de RV en las muestras de seguimiento fue realizada en 7 de los 11 pacientes, debido a la falta de disponibilidad de la secuencia genómica completa en algunos casos. Las secuencias de estos virus están depositadas en *GenBank* pendientes de asignación de identificador. De estos virus se realizó el análisis de la distancia entre pares empleando la secuencia de nucleótidos que codifica para la proteína mayoritaria de la cápside VP1 con el programa MEGA 5.2 (Materiales. Apartado 3.5.5). La diferencia de tipo de RV se establece en distancias mayores del 13% entre las secuencias de nucleótidos del gen VP1 [199,200]. De los 7 pacientes estudiados, todos los RV comparados fueron RV de tipo A excepto los RV detectados en el paciente 125 que fueron RV tipo B (Tabla 22).

Los resultados obtenidos en los pacientes 121, 122, 126 y 131 (Tabla 22) reflejaron que los RVA secuenciados fueron diferentes mes a mes. En el caso de los pacientes 127 y 128, como fue

mencionado previamente, los virus secuenciados procedían de muestras tomadas con días de diferencia por lo que por ello la homología de los virus fue del 100%.

En el caso de RVB fue observado en el paciente 125 que existía una homología del 99,3% entre el virus de la primera fecha de toma de muestras y 2 meses después confirmándose un episodio asociado al mismo tipo de RV.

	RV	RV	% de distancia en pares VP1
<b>Paciente 121</b>	RVA mes 1	RVA mes 3	20,9%
	RVA mes 3	RVA mes 5	19,8%
	RVA mes 1	RVA mes 5	20,3%
<b>Paciente 122</b>	RVA mes 1	RVA mes 2	20,7%
<b>Paciente 125</b>	RVB mes 1	RVB mes 3	1%
<b>Paciente 126</b>	RVA mes 1	RVA mes 2	21,6%
	RVA mes 2	RVA mes 3	22,3%
	RVA mes 1	RVA mes 3	24,2%
<b>Paciente 127</b>	RVA mes 1	RVA mes 1	0%
<b>Paciente 128</b>	RVA mes 1	RVA mes 1	0%
<b>Paciente 131</b>	RVA mes 1	RVA mes 8	100%

**Tabla 22:** Porcentaje de homología entre los diferentes rinovirus secuenciados por NGS de los pacientes con muestras de seguimiento

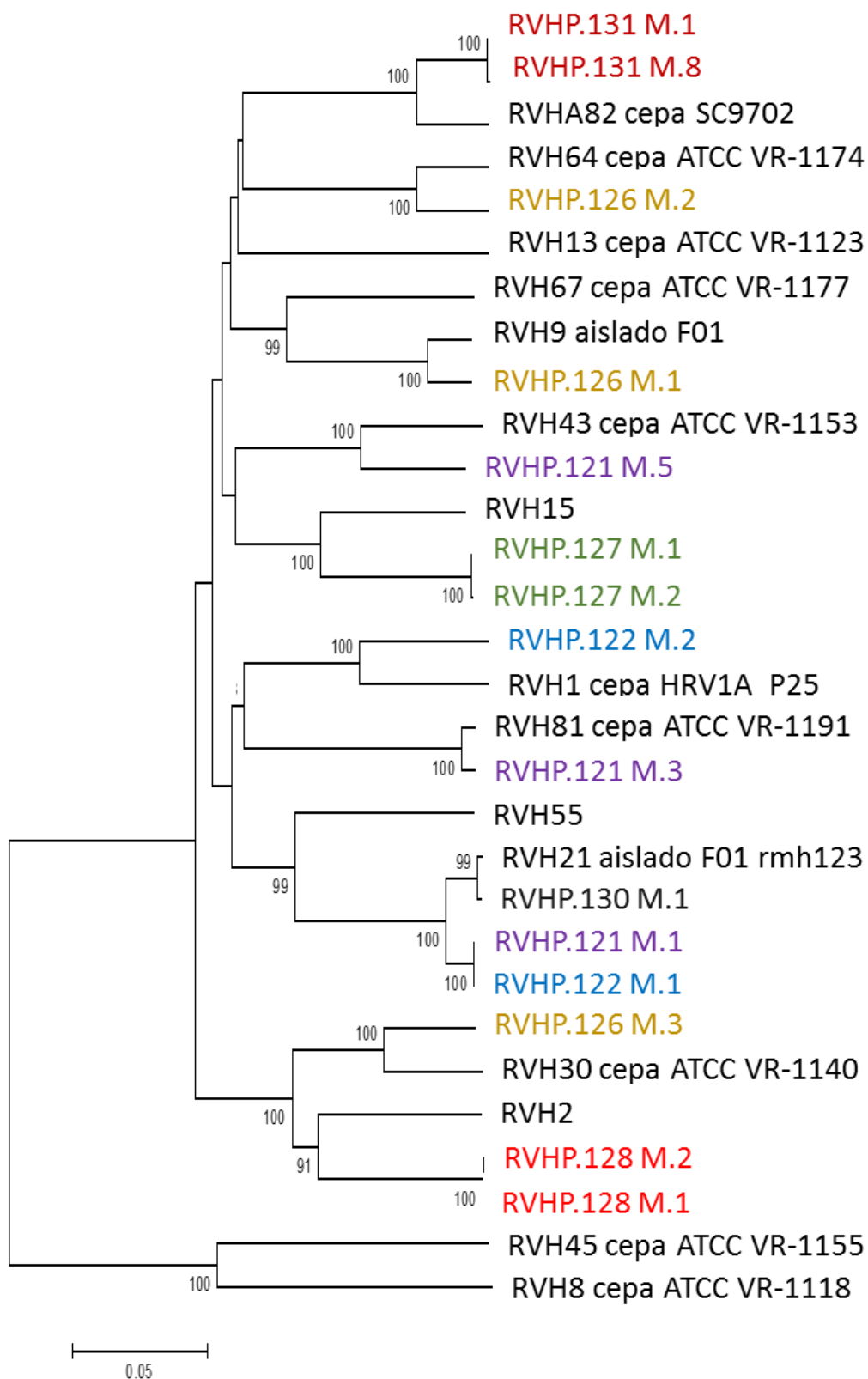
Para determinar si los virus detectados en las diferentes muestras de seguimiento se asociaban con otros rinovirus conocidos se realizó un análisis filogenético de la secuencia de nucleótidos del gen VP1 con el programa MEGA 5.2 usando el método Neighbor-Joining con un modelo Tamura-Nei, con un respaldo estadístico de 5000 repeticiones (Figura 23).

Las secuencias analizadas fueron depositadas en la base de datos *GenBank* identificándose como MN497380-MN497394.

Los RV detectados en las muestras de seguimiento no se agruparon por paciente excepto en el caso de las muestras 1 y 2 de los pacientes 127 y 128, tomadas con días de diferencia, en las cuáles se les detectó el mismo virus.

El resto de virus detectados en este apartado agrupaban con diferentes RVs de referencia demostrando la ausencia de nuevos tipos de RV y caracterizando el tipo de RV presente en cada una de las muestras de seguimiento (Figura 23). La agrupación de las muestras del paciente 131 tomadas el primer y el octavo mes y su homología del 100% puede deberse a dos procesos infecciosos independientes causados por el mismo virus.

Las diferencias en homología entre los virus detectados a lo largo del tiempo en las muestras de seguimiento y la asociación de virus detectados en las muestras del mismo paciente con diferentes tipos de RV, muestra que cada RV detectado en las muestras de seguimiento a partir de 1 mes describe un proceso infeccioso independiente del anterior.



**Figura 23:** Análisis filogenético de los gen VP1 de los RV detectados en esta tesis y rinovirus de referencia. El análisis fue realizado con el programa *MEGA 5.2* usando el método *Neighbor-Joining* con un modelo *Tamura-Nei*, con un respaldo estadístico de 5000 repeticiones. P: paciente. M: muestra.

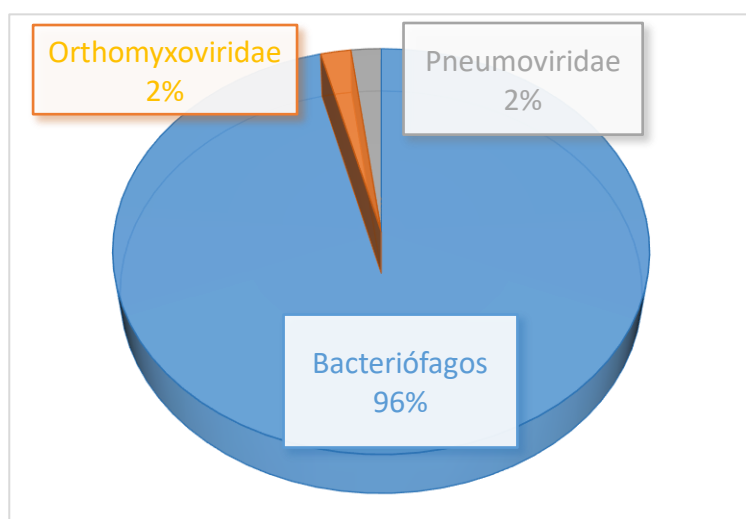
#### 5.2.1.4 Relación de RV con otros virus presentes en el tracto respiratorio

El análisis de viroma respiratorio por NGS permite la caracterización de la diversidad viral del tracto respiratorio. Los datos obtenidos de la secuenciación NGS de las muestras de seguimiento mostraron que la segunda familia de virus más numerosa tras la familia *Picornaviridae* fue la familia *Anelloviridae* (Figura 21. Página 133).

Los TTV son miembros destacados de la familia *Anelloviridae* y a pesar de que algunos trabajos les relacionen con el agravamiento de la infección de la respiratoria [54] o el desarrollo de asma [55] su papel concreto en la infección respiratoria no es conocido hasta el momento.

En el grupo control sano descrito en el Apartado 5.2.2 de este objetivo, fueron empleadas 61 ANFs procedentes de niños control sano, y 21 de ellos procedían de muestras control de seguimiento tomadas en individuos sanos (Materiales. Apartado 3.3.3).

En el análisis específico de las 21 muestras de individuos control sano con seguimiento se obtuvieron 12,6 millones de lecturas de las cuales 1246 fragmentos de genoma identificados pertenecían a secuencias virales. La identificación taxonómica de estas secuencias virales mostró que el 96% de los fragmentos pertenecían a bacteriófagos y el 4% (49,84) restante a virus que infectan a eucariotas (Figura 24). Las familias de virus que infectan a eucariotas detectadas en proporción muy baja fueron la familia *Pneumoviridae* (117 lecturas) y *Orthomyxoviridae* (53 lecturas).



**Figura 24:** Identificación taxonómica de los virus detectados tras el análisis de las muestras respiratorias de seguimiento de los individuos control sanos

Debido a la baja detección de estos anelovirus en el análisis de la composición del viroma de pacientes sanos (Figura 20. Página 131) en comparación con el viroma de pacientes con IRA (Figura 21. Página 133), se decidió emplear

el método RCA (Métodos. Apartado 4.6), por ser el más sensible en la detección de virus circulares, en las muestras respiratorias de pacientes sanos y de pacientes positivos para RV.

El análisis por RCA de las 21 muestras de seguimiento control sanos resultó negativo para TTV.

En el caso de los 11 pacientes analizados con IRA 4 de ellos fueron hospitalizados debido a la gravedad de los síntomas de la infección respiratoria que sufrían. En el análisis bioinformático de las secuencias obtenidas por NGS observamos la presencia de TTV en 4 pacientes que resultaron ser los 4 pacientes hospitalizados.

En el análisis de las muestras de pacientes con IRA positivos para RV obtuvimos 5 pacientes positivos para TTV en un total de 10 muestras respiratorias (3,5%) analizadas de estos pacientes (Tabla 23).

Paciente	Método de detección	RV en la misma muestra	Diagnóstico	Per. Hosp
Paciente 121	NGS/RCA	RVA mes 1	Sibilancias de repetición	4 días
		RVA mes 3	Sibilancias de repetición	7 días
		RVA mes 5	Sibilancias de repetición	2 días
Paciente 122	RCA	RVA mes 1	ITRS	N/A
		RVA mes 2	N/A	
Paciente 123	NGS/RCA	N/A	Sibilancias de repetición	7 días
Paciente 128	NGS/RCA	RVA mes 1	Sibilancias de repetición	7 días
		RVA mes 1	Sibilancias de repetición	
		RVC mes 3	ITRS	2 días
Paciente 130	NGS/RCA	RVB mes 5	Bronquiolitis	7 días.

**Tabla 23:** Método de detección, tipo de RV detectado y diagnóstico de los pacientes en los que fue detectado TTV

Además de los 4 pacientes hospitalizados, mediante RCA, fue detectado TTV en las 2 muestras analizadas del paciente 122.

En el caso del paciente 123, paciente hospitalizado, mediante NGS, no fue detectado el RV previamente detectado por PCR múltiple tiempo real pero si fue detectado TTV.



En el paciente 130 su hospitalización coincidió con la toma de la muestra en el mes 5, en esta muestra si fue detectado TTV a diferencia de la muestra del mes 1 en la que no fue detectado por ninguno de los métodos empleados (Tabla 23. Página 139).

Como conclusión, en nuestro estudio y con un número limitado de casos, la presencia de TTV es más alta en pacientes con infección respiratoria cuando se compara con niños sanos control. Además, en el análisis de los pacientes de seguimiento empleados en este objetivo fue detectado, mayoritariamente, este virus en pacientes hospitalizados y con infecciones recurrentes.

## DISCUSIÓN

El desarrollo de los proyectos de investigación, PI15CIII/00028 y PI18CIII/00009, coordinados con los Servicios de Pediatría del Hospital Universitario Severo Ochoa de Leganés y del Hospital Universitario La Paz de Madrid, permitió el estudio de viroma de muestras respiratorias de individuos pediátricos sanos y pacientes con IRA bien caracterizados clínicamente (Figura 16. Página 126).

El desarrollo de los métodos para el estudio de viroma respiratorio por NGS permitió el estudio de 57 ANFs de pacientes que presentaban síntomas de infección respiratoria pero a los que no se les diagnosticó ningún virus respiratorio empleando las PCRs múltiples tiempo real de caracterización virológica del laboratorio (Métodos. Apartado 4.12)

Con el fin de arrojar luz sobre la etiología de estas infecciones los 57 ANFs fueron secuenciados, junto con las 61 muestras de los individuos sanos del grupo control, empleando el flujo de trabajo para el estudio de viroma (Métodos Apartado 4.5).

El análisis de las secuencias obtenidas por NGS detectó las familias virales: *Papillomaviridae*, *Pneumoviridae*, *Picornaviridae*, *Paramyxoviridae*, *Anelloviridae*, *Coronaviridae*, *Circoviridae*, *Orthomyxoviridae* y *Polyomaviridae* (Figura 18. Página 128) y permitió la identificación de al menos un patógeno viral respiratorio en 35 de los 57 ANFs del grupo de pacientes con infección respiratoria, lo que supuso un 61% del total de los aspirados analizados. De los 57 ANFs analizados 22 (39%) de ellos permanecen sin explicación desde un punto de vista virológico, ya que no se identificó ningún virus respiratorio conocido. En el caso de estas muestras se realizó un análisis exhaustivo y en profundidad de aquellos fragmentos sin coincidencia por BLASTn

(Materiales. Apartado 3.5.5) que solo devolvió coincidencias con anelovirus, papilomavirus, virus circulares humanos similares al virus HRCiV posteriormente descrito y poliomavirus humano tipo 4.

El poliomavirus humano 4 fue detectado en 1 de las muestras con IRA analizadas pero no fue detectado en el grupo de pacientes control. Este virus se detectó originalmente en las secreciones respiratorias de un paciente pediátrico diagnosticado con neumonía de origen desconocido [201]. Estudios posteriores más amplios en series de pacientes no relacionan a este virus, poliomavirus humano 4, con la producción y el desarrollo de IRA [202]. En el paciente al que le fue detectado este poliomavirus no se le identificó ningún otro virus respiratorio, sin embargo ha sido descrito que poliomavirus humano 4 presenta una baja prevalencia en casos de infección respiratoria y frecuentemente se detecta junto con otros patógenos virales respiratorios comunes, por lo que se necesitan estudios de series de pacientes para determinar su importancia clínica [201–203].

Los anelovirus se detectan con frecuencia en la mayoría de los tejidos y órganos, incluido el tracto respiratorio de individuos sanos, y no existe asociación directa con ninguna enfermedad en humanos hasta el momento [204].

Los virus del papiloma son muy comunes en el ser humano y la mayoría de las infecciones son asintomáticas que se resuelven espontáneamente. Además de las verrugas comunes y los diversos tipos de cáncer (incluida una forma orofaríngea), asociados a la infección por el virus del papiloma humano, los tipos 6 y 11 son la causa predominante de papilomatosis respiratoria, una enfermedad en la que crecen tumores no cancerosos en las vías aéreas del tracto respiratorio [205]. Sin embargo, aparte del cáncer de orofaringe y la papilomatosis respiratoria, no hay evidencia de una asociación entre la infección por el virus del papiloma y la enfermedad respiratoria aguda. Además de la falta de evidencia de una asociación con la enfermedad respiratoria, en este estudio también se identificaron anelovirus y papilomavirus en el grupo de control, lo que refuerza la conclusión de que su presencia no estaba relacionada con la enfermedad respiratoria. En el estudio de viroma de los individuos control sanos, papilomavirus fue identificado como la familia viral con mayor número de fragmentos genómicos detectados. En concreto, el papilomavirus detectado en 11 de las muestras de los individuos control sanos fue el virus papiloma humano “Dysk 6”, que previamente ha sido descrito por NGS en un estudio de biopsias de piel y no se encuentra aún clasificado taxonómicamente [206].

El virus circular humano HRCiV se identificó recientemente en las muestras respiratorias de una paciente con IRA ingresada en la UCI que resultó negativa para los virus respiratorios analizados por las PCRs diagnósticas que se emplean en el Laboratorio de Gripe y Virus Respiratorios. Por el momento no se ha podido establecer una asociación directa con la enfermedad por lo que no está claro cuál es el valor de encontrar lecturas de secuencias que coincidan con este virus en las muestras del grupo de infección respiratoria, siendo necesarios más estudios para determinar su papel en la enfermedad respiratoria humana.

El mejor rendimiento de los métodos NGS en el estudio de estos pacientes con infección sin identificación de patógeno frente a los métodos de detección por PCR tiempo real no se debió a la presencia de patógenos desconocidos en las muestras respiratorias, sino principalmente a problemas puntuales en los métodos moleculares empleados para detectar virus respiratorios específicos. Todos los virus identificados por NGS en nuestro estudio eran patógenos respiratorios bien conocidos y no se detectaron virus nuevos.

En conjunto, los resultados obtenidos muestran que los patógenos respiratorios virales conocidos juegan un papel etiológico principal en los casos de infección respiratoria negativos. Hay que tener en cuenta que el análisis de bacterias y otro tipo de microorganismos pueden filiar la infección respiratoria en estos pacientes [207,208].

En nuestro estudio, incluimos un grupo control sano formado por una cohorte de individuos sanos. Dicho grupo control sano se comparó en todos los niveles con la cohorte de pacientes con infección con resultados negativos e infección del tracto respiratorio no filiada: 1. los individuos del grupo de control se emparejaron por edad; 2. provenían de la misma área geográfica y 3. sus muestras también se tomaron sistemáticamente durante la misma ventana de tiempo que cubre la temporada epidémica de circulación del virus. Al incluir el grupo de control, pudimos determinar que los virus detectados en las muestras respiratorias tomadas de los pacientes con IRA no circulaban entre la población sana, proporcionando evidencia de que tales virus eran responsables de la enfermedad respiratoria observada en esos casos [13]. En resumen, la inclusión del grupo de control sano nos permitió evaluar la relevancia clínica de los virus identificados en las muestras tomadas de los casos de infecciones respiratorias no explicadas.

Las familias detectadas en el análisis del grupo control fueron *Papillomaviridae*, *Pneumoviridae*, *Anelloviridae*, *Picornaviridae*, *Herpesviridae* y *Reoviridae* (Figura 20. Página 131). En este caso

solo 2 muestras contenían fragmentos atribuidos a patógenos respiratorios virales (HMNV y RV) (Tabla 20. Página 132).

Los fragmentos de HMNV detectados en una de estas muestras acumularon la gran mayoría de las fragmentos asignados a virus eucariota (99.1%) en este grupo de estudio.

La identificación de los virus respiratorios por NGS detectados se confirmó con las PCRs múltiples tiempo real BRQ y Respiratorios 2.

El HMNV y el RV identificados en el grupo control fueron detectados en 2 casos que no mostraron síntomas de infección respiratoria y que no fueron detectados por PCR tiempo real previamente. En el caso de RV las infecciones asintomáticas han sido publicadas previamente y son más frecuentes en niños pequeños [209] y aunque las infecciones asintomáticas por HMNV pueden ocurrir en la etapa pediátrica, son más frecuentes en individuos adultos inmunocompetentes [209–211].

Los hallazgos descritos en esta tesis concuerdan con estudios previos con diseño y herramientas similares [212–214]. Xu y col. [214] emplearon métodos NGS para analizar un conjunto de muestras respiratorias tomadas de niños con neumonía adquirida que arrojaron resultados negativos por métodos moleculares empleando para la detección de virus un panel respiratorio, en este trabajo fueron también detectados TTV y poliomavirus humano 4. Taboada y col. [213] estudiaron muestras de lavados nasales de niños con infecciones respiratorias que previamente resultaron negativos para patógenos respiratorios bacterianos y virales comunes al ser analizados por PCR e identificaron al menos un virus respiratorio conocido (incluidos VRS, CoV y RV) en la gran mayoría de las muestras. Ninguno de estos estudios reveló la presencia de ningún posible nuevo virus en las infecciones respiratorias no diagnosticadas analizadas y la gran mayoría de los casos podrían atribuirse a patógenos respiratorios virales conocidos.

El empleo del flujo de viroma para la secuenciación NGS nos permitió llegar a un diagnóstico específico en más de la mitad de las 57 infecciones no diagnosticadas, ya que todos los casos resueltos adicionalmente fueron producidos por patógenos respiratorios conocidos.

La identificación de lecturas virales en muestras clínicas sigue siendo controvertida porque no implica necesariamente que dichos virus sean responsables de los síntomas observados. Muchos virus pueden causar infecciones asintomáticas o subclínicas, o simplemente estar presentes en el viroma residente y replicarse sin ninguna consecuencia patogénica. Esto dificulta la

interpretación y la comprensión de los resultados obtenidos por los estudios de descubrimiento de virus realizados por NGS y relativiza la importancia clínica de los métodos de secuenciación independientes de secuencia [213].

Es conocido que los RV son uno de los principales agentes causales de infección respiratoria en pacientes pediátricos [215], en el caso de las muestras pediátricas estudiadas en el estudio de viromalos RV fueron el segundo grupo de virus más numerosos detectados en pacientes con infección respiratoria (Tabla 18. Página 128). La disponibilidad de muestras de pacientes pediátricos positivos para RV e individuos control sanos con muestra de seguimiento permitió realizar un estudio sobre la presencia de RV en muestras respiratorias del mismo paciente a lo largo de tiempo. El empleo del método de enriquecimiento con el panel de captura por sondas HyperCap Target Enrichment (Materiales. Apartado 3.5.2) permitió la obtención de tamaños de secuencia adecuados que permitían la caracterización genómica de estos virus detectados a diferentes tiempos de evolución.

Del estudio de los 35 ANFs procedentes de 11 pacientes pediátricos se obtuvieron 3,2 millones de lecturas virales de las cuales el 29% de los fragmentos genómicos pertenecían a la familia *Picornaviridae*. La distribución de los RV detectados mostró que el RVA fue el tipo mayoritario representado el 64% de los fragmentos virales pertenecientes a la familia *Picornaviridae*, seguida por RVB con un 23% y el RVC con un 13% de los fragmentos virales (Figura 22. Página 134).

Conocer la distribución de los tipos de RV presentes en el estudio de cohorte de pacientes es importante ya que el potencial patogénico de RV puede variar con el tipo de RV detectado en la muestra [216]. Estudios de cohortes de pacientes pediátricos previos han descrito generalmente mayor proporción de RVA frente a RVB y RVC [217,218]. Los resultados obtenidos en esta tesis coinciden con los trabajos previos, aunque esta representación mayoritaria de RVA y tan minoritaria de RVC puede deberse a múltiples factores como la estacionalidad anual característica de estos virus o al tipo de virus circulante en la temporada epidémica concreta. Generalmente, la temporada epidémica de RV suele establecerse al comienzo del otoño y puede presentar un pequeño pico epidémico en primavera [219], aunque trabajos previos han descrito que en el caso de RVC también pueden producirse picos de circulación epidémica en invierno en climas templados [220]. A pesar de esta posible explicación, la muestra de pacientes empleada es pequeña y futuros trabajos con una muestra de cohorte mayor son necesarios.

Los RV son considerados virus de larga persistencia determinándose la horquilla de detección de estos virus en muestra respiratoria entre 5 y 6 semanas [221]. Las muestras de seguimiento de los pacientes con IRA empleadas permitieron el estudio de RV en muestras positivas hasta 8 meses posteriores a la primera muestra analizada del paciente (Tabla 21, Página 134). De los 11 pacientes analizados se caracterizaron los RVs presentes en la muestra en 10 de ellos. De estos 10 pacientes fue posible caracterizar el seguimiento de la infección por RV en 7 de ellos.

Para determinar los diferentes tipos de las muestras analizadas se empleó el cálculo de la distancia por pares de los nucleótidos que codifican para la proteína VP1. El empleo de esta proteína se debe a que es la más externa y dominante de las proteínas de superficie picornavirales y contiene la mayoría de motivos que interactúan con los receptores celulares y los anticuerpos neutralizantes por lo que ha sido empleada en la descripción de los diferentes serotipos de los RV [222]. Se establece la diferencia de tipo de virus cuando en la secuencia de nucleótidos de VP1 hay diferencias mayores del 13% entre las secuencias [199,200]. La mayoría de virus detectados a partir de un mes presentaron un porcentaje de distancia entre pares mayor del 13% (Tabla 22. Página 135). En el caso de RVA observamos que únicamente detectamos el mismo virus en las muestras de los pacientes 127 y 128 cuyas muestras de seguimiento fueron tomadas con días de diferencia. El resto de RVA detectados en las muestras de seguimiento de los pacientes estudiados muestra diferencias en la secuencia de nucleótidos de VP1 mayores al 13%.

En el caso de RVB observamos que el virus detectado en la segunda muestra analizada del paciente 125 un mes después de la primera toma de muestra es casi idéntico al primer RVB detectado a diferencia los resultados obtenidos en la caracterización de los RVA detectados.

El análisis filogenético confirmó los resultados obtenidos en el análisis de distancia por pares. Los virus obtenidos agruparon con otros RV humanos de referencia, no agrupándose por pacientes excepto en el caso de los RV detectados en las muestras 1 y 2 de los pacientes 127 y 128. La diferencia entre los RVA detectados puede deberse a la ocurrencia de procesos infecciosos independientes causados por diferente tipo de RV en estos pacientes.

Como otros virus ARN de cadena simple en sentido positivo, los RV codifican una ARN-polimerasa ARN dependiente que carece de la capacidad de reparar errores. Las tasas de mutación estimadas para estos virus van desde  $10^{-3}$  a  $10^{-5}$  mutaciones por nucleótido por evento de replicación

[223]. Un trabajo previo estudió empleando modelos de genética de poblaciones genomas completos de RVA, RVB y RVC. Este trabajo propuso la existencia de subpoblaciones virales en muestras que presentaban poblaciones de RVA y RVC, mientras que la población de RVB era homogénea [224].

La existencia de coinfecciones con múltiples RV [44] puede conducir a la recombinación de virus con distintas propiedades biológicas y características clínicas. Se desconocen las condiciones requeridas para la recombinación de RV, pero la caracterización de RV en muestras de seguimiento de pacientes puede proporcionar un marco de referencia para este análisis [225]. Futuros estudios de metagenómica de RV podrían mostrar las ventajas de la secuenciación completa del genoma en lugar del serotipo concreto con el fin de conocer las posibles recombinaciones y mutaciones presentes en todas las regiones del genoma.

Al realizarse métodos de secuenciación independiente de secuencia pudieron ser detectadas otras familias de virus además de la familia *Picornaviridae*, las familias detectadas fueron las familias: *Anelloviridae*, *Coronaviridae*, *Pneumoviridae* y *Polyomaviridae* (Figura 21. Página 133). Como se refleja en la Figura 21 la familia con mayor porcentaje de lecturas virales tras los *Picornaviridae* fue la familia *Anelloviridae* representando el 15% de los fragmentos de virus detectados en el viroma respiratorio de estos pacientes.

La detección conjunta de TTV y RV no es la primera vez que se describe en el tracto respiratorio, la presencia de ambos virus fue previamente descrita en el viroma de pacientes adultos con IRA [56]. Como hemos mencionado previamente, la presencia del poliomavirus humano tipo 4 se ha detectado conjuntamente con otros patógenos virales respiratorios comunes [201–203].

Las coinfecciones virales de RV con otros virus respiratorios han sido detectadas en el 30% de los pacientes pediátricos con IRA por lo que la presencia de CoV y VRS en muestras con RV es biológicamente posible [226]. Trabajos previos demostraron una asociación positiva significativa entre RV y PIV, RV y VRS, y RV y Adv [227]. En el caso de los pacientes pediátricos españoles, un estudio prospectivo entre 2005 a 2013 observó que la doble infección VRS-RV producía fiebre e hipoxia en un porcentaje más alto que las infecciones por RV como único patógeno [228].

En este objetivo se realizó además el estudio de la relación de RV con otros virus presentes en el viroma respiratorio. Los TTV estos virus han sido descritos en prácticamente todo tipo de tejidos aunque suelen ser más abundantes en sangre ya que parece ser que replican fundamentalmente

en linfocitos T [229] y son muy diversos genéticamente con un espectro de virus de genomas [230] similares como torquetenominivirus y torquetenomidivirus [231].

El empleo del método RCA [14] para la amplificación del genoma de TTV fue mejor método para la caracterización de estos virus que la secuenciación por NGS.

En el estudio de las muestras control sano de seguimiento por RCA no fue detectado ningún TTV en las 21 muestras analizadas (Figura 24. Página 138), en cambio en las muestras de los pacientes con IRA fue detectado un TTV en 5 pacientes (Tabla 23. Página 139) de los cuales 4 estaban hospitalizados. En el caso del paciente 130, su hospitalización coincidió con el mes 5 de toma de muestra respiratoria, fue en esta muestra, durante su hospitalización, en la que TTV fue detectado y no en las muestras previas.

De acuerdo al papel de este virus en el tracto respiratorio Maggi y col. [232] propusieron que estos virus solos o en sinergia con otros virus, podrían actuar como potenciadores de la inflamación en el tracto respiratorio superior e inferior. En pacientes pediátricos con IRA se caracterizó el tracto respiratorio superior como el lugar de infección primaria y replicación de TTV [54]. Además, se ha demostrado la alta prevalencia de infecciones de TTV en niños con graves afecciones respiratorias estableciéndose correlación entre la proporción de TTV presente en el tracto respiratorio y la limitación de función del tracto respiratorio periférico [55].

A pesar de estos estudios sobre TTV, infección respiratoria y relación con el sistema inmune el rol concreto de este virus en el desarrollo de IRA o asma continúa por ser descrito. A pesar de ser detectado fundamentalmente en pacientes con IRA grave también fueron detectados en el estudio de viroma de los individuos control sanos (Figura 20. Página 131). La falta de detección en las muestras de control sano de seguimiento analizadas por RCA puede deberse a que el tamaño de muestra no es suficiente por lo que estudios futuros son imprescindibles para entender la influencia de estos virus en la infección respiratoria o el desarrollo de asma.

El empleo del flujo de trabajo para el estudio de viroma aplicado a las muestras respiratorias permitió observar una foto de conjunto de los virus presentes en salud y enfermedad abriendo la puerta a trabajos futuros sobre subpoblaciones virales de virus como RV, sobre las relaciones ecológicas entre los virus presentes durante la infección respiratoria y sobre la comprensión de la diversidad viral del tracto respiratorio.



## 5.2.2 Línea de trabajo 2: El estudio de viroma como herramienta para descubrir nuevos virus

Un nuevo virus CRESS-DNA detectado en dos pacientes con infección respiratoria grave que resultaron negativos para otros patógenos respiratorios es descrito en este objetivo. Este virus ha sido llamado *Human Respiratory Circular Virus* (HRCiV) haciendo referencia a su estructura y al hospedador donde fue detectado.

La detección del virus se realizó gracias al empleo del flujo de trabajo para el estudio de viroma (Métodos. Apartado 4.5). La preparación de librerías para secuenciación por NGS se realizó usando el kit de preparación de librerías *Nextera XT* (Materiales. Apartado 3.5.2) con el secuenciador *MiSeq* (Materiales. Apartado 3.5.4). El análisis bioinformático de las muestras fue realizado como se describe en el Apartado 4.10 de Métodos empleando la plataforma *Virus TAP* [24] y las herramientas de análisis empleadas en USAMRIID.

Para el análisis de la presencia del virus HRCiV en el tracto respiratorio se realizó un estudio en el que se emplearon 3 grupos: pacientes adultos con infección respiratoria (Grupo 1); pacientes pediátricos con infección respiratoria (Grupo 2), e individuos control sanos (Grupo 3). Las características de las muestras respiratorias empleadas en este estudio fueron descritas en el Apartado 3.3.3.3 de Materiales (Figura 3. Página 69).

En el Grupo 1 se analizaron 368 muestras respiratorias de mayores de 16 años con infección respiratoria enviadas por hospitales y centros de salud para el diagnóstico de gripe dentro del Sistema de Vigilancia de la Gripe en España. En el Grupo 2 fueron analizadas 361 muestras respiratorias de pacientes pediátricos con infección respiratoria procedentes de los proyectos de investigación PI15CIII/00028 y PI18CIII/00009.

El Grupo 3 de pacientes sanos fueron empleados como grupo control del estudio, este grupo consistió en 61 muestras tomadas a pacientes pediátricos que no presentaron cuadro clínico compatible con infección respiratoria y que fueron previamente empleadas en el estudio de viroma respiratorio en la “Línea de trabajo 1” de este objetivo. Previamente al estudio de presencia de virus en el tracto respiratorio de los pacientes de los diferentes grupos fueron realizadas las PCRs de diagnóstico de virus tiempo real múltiples BRQ y Respiratorios 2 y Gripe A, B y C (Método. Apartado 4.12. Páginas 91-93)

El empleo de la extracción manual por precipitación alcohólica de los ácidos nucleicos totales

(Métodos. Apartado 4.4.1) fue el protocolo empleado para descartar contaminaciones ambientales de las membranas de las columnas que pudieran producir resultados falsos positivos. El diseño de una PCR de detección para este nuevo virus permitió el hallazgo de otros nuevos virus filogenéticamente emparentados.

Para la caracterización de los virus detectados se combinó el empleo del método de amplificación de virus circulares RCA (Métodos. Apartado 4.6) y de la PCR convencional de genoma completo para la obtención del genoma completo de los nuevos virus.

Aunque estos virus presentan las características comunes de los miembros de la familia *Circoviridae* [233], el tamaño de su genoma y la baja homología genética con otros miembros conocidos de esta familia sugirió que podrían formar parte de un grupo de virus dentro de la recientemente propuesta familia *Redondoviridae* [99]

#### **5.2.2.1 Caracterización virológica de pacientes con neumonía grave**

La paciente 1 de este objetivo fue una mujer de 69 años ingresada en la UCI del Hospital Universitario Puerta de Hierro (Majadahonda, Madrid) con fiebre superior a 38° C, tos, rinorrea, astenia, dificultad para respirar, dolor abdominal y diarrea (Materiales. Apartado 3.3.3.1). Su historial de viaje indicaba una visita a Oriente Medio por lo que se activó el protocolo de alerta ante la sospecha de infección por CoV- MERS.

En total, 12 muestras fueron tomadas durante los 8 días de ingreso en UCI y enviadas al laboratorio de Referencia de Gripe y Virus Respiratorios (Tabla 5. Página 68). Estas muestras fueron analizadas por PCR múltiple tiempo real de diagnóstico para virus respiratorios (Métodos. Apartado 4.12), incluyendo el coronavirus MERS, resultando todas ellas negativas para todos los virus analizados (Tabla 24).

A causa de la gravedad de la paciente y el resultado negativo de las pruebas previas de diagnóstico fue adoptada una estrategia diferente: se realizó el análisis por NGS de las 12 muestras respiratorias de la paciente 1.

Estas muestras fueron tratadas siguiendo el flujo de trabajo para la realización del estudio de viroma (Métodos. Apartado 4.5) y secuenciadas con la plataforma *MiSeq*.

El análisis bioinformático mostró la obtención de un total de 252.345 lecturas virales. El 84% de los fragmentos virales incluían a bacteriófagos de la familia *Myoviridae*, *Podoviridae*, *Siphoviridae*

y *Microviridae*. El 16% restante correspondían con virus herpes (34.885 lecturas), el virus “Human respiratory-associated PSCV-5-like” (HRAPLV) (5.435 lecturas), picobirnavirus animales y humanos (432 lecturas) y anellovirus (48 lecturas).

Los resultados obtenidos fueron llamativos debido a la ausencia de patógenos respiratorios conocidos y la presencia en 5 de las 12 muestras de secuencias coincidentes con el virus HRAPLV (GenBank: KY052047) (Tabla 24).

Identificación de muestras y días desde la admisión en el hospital												
	Día 1			Día 2			Día 3		Día 4		Día 7	Día 8
NGS	15016	15017	15025	15027	15028	15039	15040	15065	15066	15067	15078	15093
Bacteriófagos	0	587	56334	51412	30898	5	46152	3	2	0	26191	0
Herpesvirus	0	0	72	8	0	0	61	0	0	0	34714	0
Picobirnavirus	0	0	0	87	336	0	0	0	0	0	0	0
Anellovirus	0	0	0	0	0	0	17	6	2	0	23	0
HRPLAV	0	0	1833	13	3	0	1688	0	0	0	1898	0
PCR HRCiV REP	-	-	+	+	-	-	+	-	-	-	+	-
Acceso GenBank	N/A	N/A	KY579361	KY579360	N/A	N/A	KY244146	N/A	N/A	N/A	KY579362	N/A

**Tabla 24:** Número de lecturas virales de los extractos ADN y ARN tras el análisis por secuenciación masiva y los resultados obtenidos en la PCR de estudio de las muestras clínicas de la paciente 1.

Con respecto al tamaño de secuencia obtenido del virus coincidente con el virus HRAPLV (KY052047) en el análisis por NGS de las muestras de la paciente 1 se obtuvo en 3 de las 5 muestras el virus casi completo. En las muestras 15025 (ENF del día 2), 15040 (ENF día 3) y 15078 (ENF día 7) pudieron ser ensamblados fragmentos largos que conformaron secuencias de 2769 pares de bases (1833 lecturas), 2875 pares de bases (1688 lecturas) y 1460 pares de bases (1898 lecturas) respectivamente.

Tiempo después de este análisis en el que se detectó este nuevo virus circular se recibieron en el laboratorio de Referencia de Gripe y Virus Respiratorios 3 muestras respiratorias (Tabla 6. Página 69) procedentes de un segundo paciente (Materiales. Apartado 3.3.4.2).

El paciente 2 fue ingresado en la UCI de un hospital terciario de Madrid presentando fiebre, tos, rinorrea y dificultad para respirar. Su historial de viaje indicaba una visita a Oriente Medio por lo que en este caso también se activó el protocolo de alerta ante la sospecha de infección por CoV-MERS.

Como en el caso anterior las muestras respiratorias del paciente 2 fueron analizadas por la batería de PCRs múltiples tiempo real de diagnóstico de virus (Métodos. Apartado 4.12), incluyendo el coronavirus MERS, resultando en este caso negativas para todos los virus analizados nuevamente.

Identificación de la muestra/ Número de lecturas		
	162311 ENF	162355 LBA
HRAPLV	0	4926 extracto ADN 2023 extracto ARN
PCR HRCiV 2	+	+
Número de acceso en <i>Genbank</i>	N/A	MK956926

Tabla 25: Número de lecturas coincidentes con el virus HRPLAV obtenidas por secuenciación masiva y el resultado de la PCR de estudio realizada a las muestras respiratorias del paciente 2.

El análisis bioinformático de las secuencias obtenidas por NGS del LBA (162355) solo detectó una coincidencia viral que también correspondía con el virus HRPLAV (KY052047) (Tabla 25). Fueron obtenidas un total de 4926 correspondían al circovirus mencionado previamente. El tamaño de secuencia obtenido fue un fragmento largo de 2526 pb en el análisis del viroma ADN.

El análisis de viroma ARN se realizó con el fin de estudiar si era posible la detección de ARN en las muestras de estos pacientes, en el caso de la muestra 162355 se obtuvo un fragmento de 672 pares de bases con 2023 lecturas coincidentes con el virus HRPLAV (KY052047) en el análisis de viroma ARN (Tabla 25).

#### 5.2.2.2 Secuenciación de virus circulares. RCA vs SISPA

Con el fin de descartar que la detección de los virus CRESS-DNA fuera fruto de una contaminación por el empleo de columnas de sílice en la extracción de ácidos nucleicos se realizó una extracción manual por precipitación alcohólica de las muestras procedentes de los pacientes 1 y 2 (Métodos. Apartado 4.4.1).

En este caso se emplearon dos métodos de amplificación de virus: el método SISPA ADN con enriquecimiento por 3 métodos como se describe en el flujo de trabajo de viroma y la amplificación por el método RCA, específica para virus circulares (Métodos. Apartado 4.6).

Las librerías de los productos de amplificación RCA se prepararon con el kit *Nextera XT* (Materiales. Apartado 3.5.2) y fueron secuenciadas empleando el secuenciador *MiSeq* (Materiales. Apartado 3.5.4).

Los resultados obtenidos tras el análisis bioinformático mostraron que existían diferencias significativas en los resultados obtenidos según el método de amplificación aleatoria empleado. Empleando la RCA fue posible la detección y la secuenciación de virus completo (Tabla 26) en 5 de los 7 virus detectados empleando este método. Las secuencias completas de los virus detectados fueron depositadas en *GenBank* como está reflejado en las Tablas 25 y 26.

Identificación	SISPA ADN/N° de pb	RCA/N° de pb
15016	-	-
15017	-	-
15025	+ / 2769 pb	<b>+ / 3054 pb</b>
15027	+ / 546 pb	<b>+ / 3054 pb</b>
15028	+ / 123 pb	-
15039	-	-
15040	+ / 1688 pb	<b>+ / 3054 pb</b>
15065	-	-
15066	-	-
15067	-	-
15078	+ / 1460 pb	<b>+ / 3054 pb</b>
15093	-	-
162311	-	-
162355	+ / 2526 pb	<b>+ / 3013 pb</b>

**Tabla 26:** Identificación, detección y tamaño de secuencia obtenidos por los métodos SISPA ADN y RCA.

Empleando el método SISPA fueron detectados los nuevos virus CRESS-DNA en 7 de las muestras respiratorias analizadas pero en ningún caso se pudo obtener secuencia completa de estos virus circulares. Además el empleo de este método no es válido para demostrar la circularidad de este nuevo virus.

Debido a los resultados obtenidos el método empleado en esta tesis para la amplificación de virus circulares fue el método RCA dada su mayor sensibilidad y la obtención de un mayor tamaño de secuencias de los virus presentes en las muestras analizadas.

### 5.2.2.3 Caracterización del virus CRESS-DNA

Con el fin de caracterizar los CRESS DNA virus hallados tanto en el paciente 1 “caso índice” como en el paciente 2, se realizó el método RCA que demostró la circularidad del virus y permitió la obtención del genoma completo.

La obtención del genoma completo del nuevo CRESS DNA virus en varias muestras de los pacientes confirmó la presencia de genoma funcional y permitió conocer las características del genoma de estos nuevos virus CRESS-DNA.

El tamaño obtenido de la secuencia completa del nuevo CRESS DNA virus fue de 3.054 nucleótidos en los virus detectados en la paciente 1 y de 3.013 nucleótidos en el virus detectado en el paciente 2.

El estudio de la secuencia del genoma completo del nuevo CRESS DNA virus usando el buscador de marcos de lectura *ORF Finder* (Materiales. Apartado 3.5.5) mostró la existencia de tres marcos de lectura, dos de ellos en sentido positivo y otro en sentido negativo. Estos resultados demostraron que el nuevo CRESS DNA virus presentaba un genoma con el sentido bidireccional característico de los circovirus.

El primer marco de lectura presentó un tamaño de 1593 nucleótidos y se transcribía en sentido negativo. Tras realizar un análisis con la plataforma *BLASTx* (Materiales. Apartado 3.5.5) de proteínas fue identificado un fragmento de 530 aminoácidos que presentaban una elevada homología con la proteína de la CAP de circovirus. La proteína CAP con más homología con la que fue asociada fue la del virus HRPLAV (GenBank: AQW45175) presentando un 77,78% de identidad con una cobertura del 100%.

El marco de lectura 2 tenía un tamaño de 1047 nt, 348 aa, y se transcribía en sentido positivo. Estos 348 aa codificaban para la proteína REP. El análisis mediante *BLASTx* de proteínas mostró que la coincidencia más próxima fue la proteína REP del virus “Porcine stool-associated circular virus 5” (PoSCV5) mostrando un 53,80% de identidad y 97% de cobertura (GenBank: YP\_009021877).

Tras el análisis de la secuencia de aminoácidos fueron localizados en la proteína REP la presencia de los motivos conservados relacionados con la replicación en círculo rodante característicos de la familia *Circoviridae*. Los motivos RCR I, II y III (FATYL, PHLER y YCQK) se encontraron en el

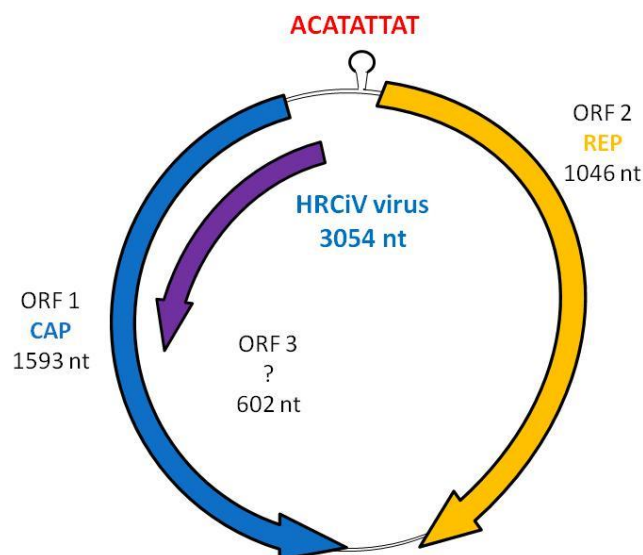
extremo N-terminal en las posiciones 12-16, 85-89 y 100-103 respectivamente. Los motivos Walker A, B y el motivo C (GETGEGKT, IIDDF e ITSP) estaban presentes en las posiciones 76-183, 218-222 y 260-263 respectivamente.

Un tercer marco de lectura en sentido negativo de 603 nucleótidos de tamaño fue obtenido, este marco de lectura codifica para una proteína hipotética de función desconocida de 200 aminoácidos, este tercer marco de lectura es habitual encontrarlo en los virus de la familia *Circoviridae*.

El virus presentaba dos regiones intergénicas con respecto a las dos proteínas codificantes con función conocida. La región intergénica 1 se encuentra entre los extremo 5' del marco de lectura 1 y 2. Esta región presenta un tamaño de 175 nucleótidos y contiene una de las principales características de los circovirus: la estructura en bucle con el nonámero conservado ACATTATAT donde se encuentra el origen de replicación.

La región intergénica 2 posee un tamaño de 239 nucleótidos y se encuentra entre los extremos 3' de los marcos de lectura 1 y 2.

Este nuevo virus fue llamado *Human Respiratory Circular Virus* (HRCiV) reflejando sus características como virus circular y el hospedador donde fue detectado (Figura 25).



**Figura 25:** Representación gráfica del genoma circular del virus HRCiV

#### 5.2.2.4 Explorando la presencia del HRCiV en el tracto respiratorio

Con el fin de realizar un estudio sobre la presencia del virus en el tracto respiratorio se diseñó una PCR convencional denominada REP HRCiV (Métodos. Apartado 4.13.1) para la detección del nuevo virus CRESS-DNA en las muestras respiratorias de 3 grupos de pacientes (Materiales. Apartado 3.3.3.).

Identificación	Diagnóstico	Presencia de otros virus
<b>181449</b>	<b>Neumonía doble</b>	<b>NA</b>
<b>181554</b>	Cuadro Gripal	Gripe A
<b>181539</b>	Cuadro Gripal	Gripe A
<b>181535</b>	Cuadro gripal	Gripe B
<b>181562</b>	Cuadro gripal	Gripe A
<b>181701</b>	Cuadro gripal grave con hospitalización	Gripe B
<b>181694</b>	Cuadro gripal grave con hospitalización	Gripe B
<b>181690</b>	Cuadro gripal grave con hospitalización	Gripe B
<b>181793</b>	Cuadro Gripal	Gripe B
<b>181752</b>	<b>Neumonía lobular atípica</b>	<b>N/A</b>
<b>181753</b>	Cuadro Gripal	Gripe A
<b>181810</b>	Cuadro Gripal	Gripe B
<b>181807</b>	Cuadro Gripal	Gripe A
<b>181796</b>	Cuadro Gripal	Gripe B
<b>181798</b>	Cuadro Gripal	Gripe A
<b>181787</b>	Cuadro Gripal	Gripe B
<b>181968</b>	Cuadro Gripal	Gripe B
<b>181960</b>	Cuadro Gripal grave con hospitalización	Gripe A
<b>182096</b>	Resfriado	RV
<b>182064</b>	Cuadro Gripal	Gripe B
<b>182377</b>	<i>Exitus</i>	Gripe B
<b>182320</b>	<i>Exitus</i>	Gripe B
<b>182276</b>	Resfriado	CoV
<b>180696</b>	Cuadro Gripal	Gripe B
<b>162298</b>	<b>Neumonía doble</b>	<b>N/A</b>
<b>15278</b>	Cuadro Gripal	Gripe A
<b>15037</b>	Cuadro Gripal	Gripe A
<b>15232</b>	Cuadro gripal	Gripe A
<b>15307</b>	<b>Neumonía grave</b>	<b>N/A</b>

**Tabla 27:** Identificación, diagnóstico y virus respiratorios encontrados en los pacientes positivos a HRCiV del grupo 1. Los pacientes que solo presentaron infección con HRCiV están destacados en negrita

Tras la realización de la PCR convencional REP HRCiV en las 368 muestras respiratorias analizadas del Grupo 1 resultaron positivas 30 de ellas (8,2%). Dentro de estas 30 muestras a las que se les detectó el virus HRCiV destacaron las de 4 pacientes que resultaron negativos para todos los virus



respiratorios analizados y además compartían el diagnóstico clínico de neumonía grave (Tabla 27).

En el de Grupo 2 tras el análisis de 361 muestras respiratorias de pacientes pediátricos se obtuvieron únicamente 3 muestras positivas (0,8%), en estas muestras respiratorias positivas fueron detectados otros virus respiratorios (Tabla 28).

Identificación de la muestra	Diagnóstico	Presencia de otros virus respiratorios
<b>SO 14627</b>	Bronquiolitis	RV
<b>SO 14760</b>	Sibilancias de repetición	Gripe A
<b>SO 14863</b>	Sibilancias de repetición	AdV, EV68

**Tabla 28:** Diagnóstico y presencia de otros virus en los pacientes positivos del grupo 2.

El Grupo 3 de pacientes consistió en 61 muestras tomadas a pacientes pediátricos que no presentaron cuadro clínico compatible con infección respiratoria. Tras el análisis de estos pacientes mediante PCR convencional REP HRCiV no se obtuvieron resultados positivos del nuevo HRCiV.

Una de las ventajas de este estudio, además del gran número de muestras disponibles para analizar, es la documentación clínica disponible de todas ellas. El estudio del seguimiento clínico de los pacientes pediátricos reveló la disponibilidad de una muestra correspondiente a un paciente de 2 años de edad cuya fecha de toma fue previa a la del estudio realizado (Hermano 1) y además de este mismo paciente se disponía de una muestra tomada el mismo día a su hermano (Hermano 2) (Tabla 29).

El resultado obtenido tras el análisis mediante la PCR REP HRCiV de la muestra previa del Hermano 1 (SO 13831) fue negativo para la detección de HRCiV. La muestra del Hermano 2 también resultó ser negativa para HRCiV tras el análisis por REP HRCiV (Tabla 29).

	Hermano 1		Hermano 2
	SO 13831	SO 14760	SO 14758
<b>Edad</b>	3 meses	3 años	2 años
<b>Fecha toma de muestra</b>	14/01/2016	14/03/2018	14/03/2018
<b>PCR HRCiV 2</b>	-	+	-

**Tabla 29:** Fecha de toma de muestra, edad y resultados de la PCR REP HRCiV de los hermanos pediátricos de estudio.

Podemos concluir que el virus HRCiV es más común en el tracto respiratorio de adultos que en niños ( $p$ -valor < 0,01). Este virus se encuentra ausente en muestras pediátricas de pacientes sanos.

#### 5.2.2.5. Papel de los virus HRCiV en la nueva familia *Redondoviridae*

Con el fin de caracterizar la distancia genética entre los virus detectados en nuestro estudio de caracterización y los virus pertenecientes a la recientemente publicada familia *Redondoviridae* [99], fue realizado un análisis de distancia entre pares y un análisis filogenético de estos virus tras la realización de un alineamiento múltiple de secuencias.

Para la realización de estos análisis se obtuvo la secuencia completa de algunos de los virus detectados en el estudio de previo de caracterización empleando la PCR convencional de genoma completo HRCiV (Métodos. Apartado 4.13.2).

En el análisis inicial fue comprobada la distancia en pares de la secuencia de nucleótidos del genoma completo entre los virus HRCiV obtenidos en este trabajo empleando el programa *MEGA* 5.2 (Materiales. Apartado 3.5.5). Los resultados obtenidos mostraron la presencia dos tipos de virus diferentes: el virus HRCiV 15278 y el resto de virus que eran muy similares al virus HRCiV 15040 (Tabla 30). La distancia en pares de los virus HRCiV 15040 y HRCiV 15278 fue de 24,4%.

Virus HRCiV del estudio de caracterización	HRCiV paciente 1	% de distancia en pares
<b>162355</b>	15040	1.8
<b>15037</b>	15040	2
<b>15307</b>	15040	1
<b>15232</b>	15040	6
<b>15278</b>	15040	24.4

**Tabla 30:** Porcentaje de distancia en pares entre los virus encontrados en el estudio de caracterización y la paciente 1

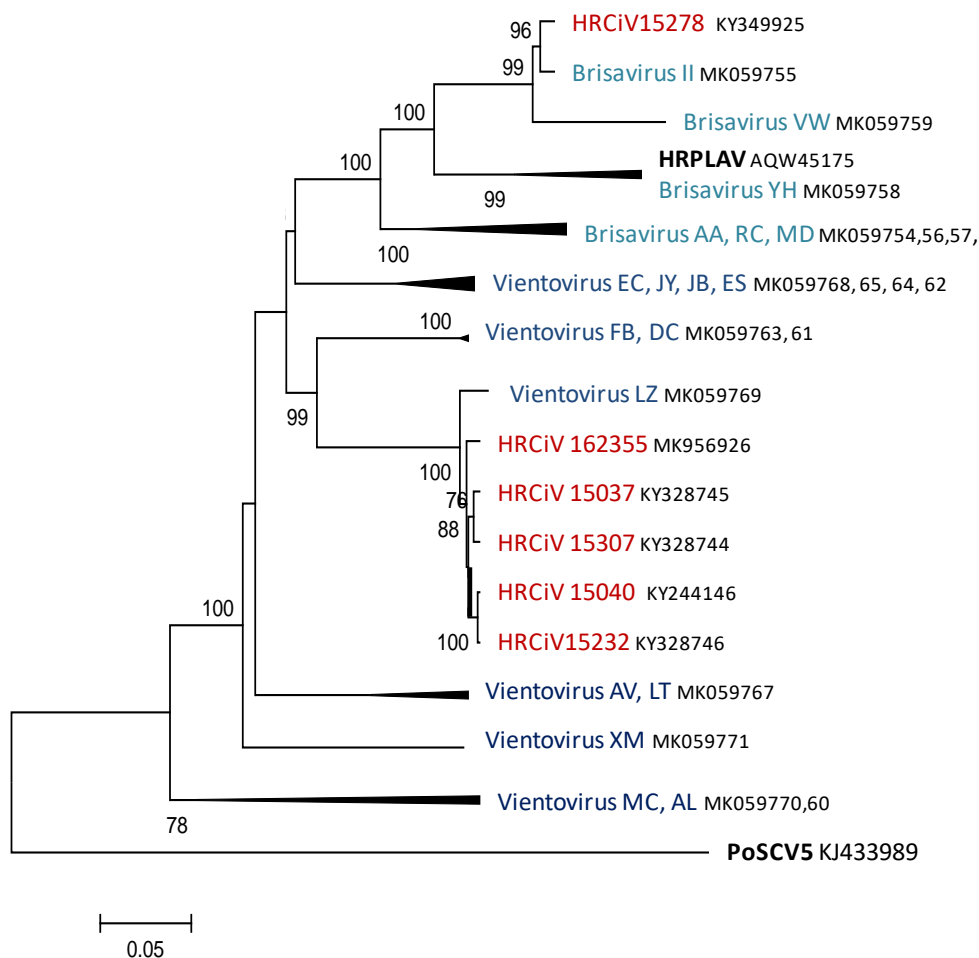
De acuerdo a las estructuras genéticas del virus HRCiV 15278 observamos que presenta motivos conservados RCR y Walker de la proteína REP diferentes de los del virus HRCiV15040. Los motivos RCR de HRCiV15278 caracterizados I, II y III fueron MLTYND, PNIES y YVIK respectivamente, y los Walker A, B, y C localizados fueron GATGSGKT, VLDDL y ITCP (Figura 26).

HRCiV15040	HRCiV 15278
RCR I <b>FATYL</b>	RCR I <b>MLTYND</b>
RCR II <b>PHLER</b>	RCR II <b>PNIES</b>
RCR III <b>YCQK</b>	RCR III <b>YVIK</b>
Walker A <b>GETGEGKT</b>	Walker A <b>GATGSGKT</b>
Walker B <b>IIDDF</b>	Walker B <b>VLDDL</b>
Walker C <b>ITSP</b>	Walker C <b>ITCP</b>

**Figura 26:** Motivos conservados RCR y Walker de la proteína REP de los virus HRCiV 15040 y HRCiV 15278

Tras la reciente publicación de la familia *Redondoviridae* fue realizado el análisis de distancia entre pares con los virus HRCiV detectados en el estudio de caracterización y los miembros de esta familia publicados. Los resultados obtenidos mostraron que el virus HRCiV 15040 presentaba una distancia de 3,17 % con el virus “*Human oral-associated videntovirus LZ*” (MK059769) y el virus HRCiV 15278 presentaba una distancia del 2% con el virus “*Human lung-associated brisavirus II*” (MK059755).

Los resultados obtenidos en el análisis de distancia en pares se reflejan en el análisis filogenético realizado con la secuencia de nucleótidos de los genes CAP y REP y concatenados empleando el método de análisis *Maximum Likelihood* con el modelo *Hasegawa-Kishino-Yano*, aplicando un respaldo estadístico de 5000 repeticiones realizado con el programa *MEGA 5.2* (Materiales. Apartado 3.5.5). Este análisis reflejó la presencia de dos tipos de virus HRCiV detectados en el estudio de caracterización (HRCiV 15040 y HRCiV 15278). Como fue representado en el primer análisis se observó que el virus HRCiV 15040 y el resto de virus detectados similares se agrupaban con el virus “*Human oral-associated videntovirus LZ*” (MK059769) del género videntovirus y el virus HRCiV 15278 agrupaba con el virus “*Human lung-associated brisavirus II*” (MK059755) mostrando la representación de ambos géneros de la familia *Redondoviridae* en los virus detectados en esta tesis (Figura 27).



**Figura 27:** Análisis filogenético de los virus HRCiV detectados en esta tesis y los virus de la familia *Redondoviridae* publicados hasta el momento. El análisis fue realizado con el programa *MEGA 5.2* usando el método *Maximum Likelihood* con un modelo *Hasegawa-Kishino-Yano*, con un respaldo estadístico de 5000 repeticiones.

#### 5.2.2.6 Aislamiento viral del virus HRCiV en diferentes líneas celulares

Con el fin de demostrar la replicación de este nuevo virus HRCiV en células se realizaron diferentes ensayos para poder aislar el virus en cultivos celulares.

Se realizó una batería de infecciones en las células VERO E-6, L132, Hep2 y MDCK-SIAT (Materiales. Apartado 3.1) que tenían como fin la prueba de diferentes líneas celulares para el aislamiento del virus HRCiV. Estas células fueron infectadas siguiendo el protocolo propuesto en el Apartado 4.2 de Métodos, el medio post-infección empleado estaba compuesto por DMEM suplementado con 1% de antibiótico y glutamina y 2% de SFB (Métodos. Apartado 4.2). Finalmente no fue posible el aislamiento viral en ninguna de estas líneas celulares.

El segundo abordaje estableció una prueba con diferentes condiciones en la composición del medio post-infección empleando la línea celular L-132. Los diferentes medios se componían de medio DMEM con suero y sin suero; con suero y 0,2% de tripsina; sin suero y 0,2% de tripsina; con suero y 1% de tripsina y sin suero y 1% de tripsina (Métodos. Apartado 4.2).

En ningún caso se obtuvo un aislamiento positivo de virus tras la infección de las respectivas líneas celulares con los diferentes tratamientos en ninguno de los abordajes.

## DISCUSIÓN

La llegada al laboratorio de Gripe y Virus Respiratorios de 12 muestras de una paciente grave ingresada en UCI a la que se le realizó el estudio de viroma al ser descartada la presencia de RV, EV, VRS, BoV, HMNV, AdV, Gripe A, B y C, VPI y CoV (incluyendo MERS-CoV) como los patógenos causantes de la infección permitió el hallazgo de un nuevo virus CRESS-DNA.

Trabajos previos empleando metagenómica habían relacionado este grupo de virus con infección respiratoria [85,86]. A pesar de formar parte del grupo de los virus CRESS-DNA, estos virus presentaban baja homología entre sí y provenían de casos clínicos con perfiles diferentes por lo que el estudio y caracterización de este nuevo virus cobraba importancia.

Centrándonos en el análisis de las secuencias obtenidas en la paciente 1 por NGS observamos que fueron detectados bacteriófagos, virus herpes humanos, picobirnavirus y anelovirus. La presencia de estos virus podría responder a procesos de infección independientes que no tendrían impacto en la infección respiratoria.

Los anelovirus han sido encontrados en todo tipo de tejidos de casi toda la población mundial independientemente de la edad, situación socio-económica y estado de salud [232], a pesar de su ubicuidad no se le conoce efecto patogénico específico en infección respiratoria. Los picobirnavirus son virus que han sido descritos en las heces de un amplio rango de hospedadores. En el ser humano su patogenicidad no queda clara al haber sido detectados en muestras de heces normales y heces procedentes de diarrea [234]. En el estudio de viroma de la paciente 1, los picobirnavirus solo fueron encontrados en dos muestras y la mayoría de lecturas provinieron de la muestra de heces 15028 (Tabla 24). Los síntomas iniciales refieren que la paciente antes de la admisión al hospital tenía diarrea por lo que podría explicarse la presencia de los picobirnavirus

en dichas muestras. Plantea más dudas el hallazgo en el esputo 15027 de estos virus, al ser la muestra anterior pudo tratarse de una contaminación cruzada que no pudo ser determinada al no quedar material disponible.

Uno de los resultados más destacados al ser un patógeno conocido es la aparición en la muestra 15078 de 34.714 lecturas de un virus herpes. Este resultado fue confirmado mediante cultivo celular y visualización de los virus mediante microscopía electrónica. Los virus herpes humanos son bien conocidos por su latencia tras la infección inicial y su reactivación puede producirse en circunstancias como la inmunosupresión, trauma local o estrés sistémico [235]. Considerando todas las complicaciones clínicas que sufrió la paciente durante su estancia en la UCI (Materiales. Apartado 3.3.3.1), la presencia del herpes pudo deberse a una recurrencia herpética.

Dada la severidad y gravedad de la infección y teniendo en cuenta que el único virus que se repite mayoritariamente en las muestras de la paciente 1 (Tabla 24. Página 151) y que en el paciente 2 (Tabla 25. Página 152) es el único virus detectado junto con bacteriófagos, y fue encontrado tanto en viroma ADN y ARN, este nuevo virus CRESS-DNA podría ser responsable de la situación clínica de los pacientes.

Trabajos previos sobre nuevos virus descubiertos por NGS mostraron la posibilidad de contaminación ambiental debido al uso de columnas de sílice durante la extracción [236] por ello para demostrar que este CRESS-DNA virus no era fruto de una contaminación se realizó la extracción mediante precipitación alcohólica de los ácidos nucleicos totales [161] y un estudio de niños control sano de 61 muestras. La ausencia total de virus en los análisis realizados por PCR HRCiV 2 y secuenciación por NGS de los pacientes control sin el uso de columnas descarta la posibilidad de contaminación.

La técnica RCA es muy popular en el estudio de virus circulares por NGS gracias a que su carácter isotérmico evita que los pasos de desnaturalización de la PCR puedan dañar estos virus [15], además al no tener cebadores dirigidos a ningún virus puede detectar diferentes poblaciones de virus [237] e incluso virus desconocidos.

Trabajos previos compararon la cobertura obtenida por RCA y SISPA medida en tres indicadores de cobertura (curvas de Lorenz, coeficiente de variación de cobertura y correlación de Pearson con la cobertura de viroma sin amplificar) de genomas únicos mostrando a la RCA como mejor método para obtener mejores coberturas y tamaño de secuencia que la amplificación por SISPA

[192].

Los resultados obtenidos en la comparación del estudio del virus circular HRCiV por RCA y SISPA corroboran lo descrito previamente pudiendo ser secuenciados 5 de los 6 virus previamente detectados obteniéndose el genoma completo en todos ellos (Tabla 26. Página 153).

A pesar de que la RCA es una técnica más sensible para la detección de CRESS DNA virus, este método solo puede ser empleado en el caso de que el objetivo sea la secuenciación de virus que permitan la amplificación por círculo rodante por lo que no podría ser utilizado en el estudio de viroma respiratorio.

La obtención de la secuencia completa del virus HRCiV 15040 disponible en la plataforma *GenBank* desde octubre de 2018 mostró como virus más homólogo al virus HRPLAV (KY052047) publicado en noviembre de 2017. A pesar de que posteriores análisis de distancias han revelado que podría formar parte de la familia *Redondoviridae* la publicación de estos virus en mayo de 2019 hace que los análisis bioinformáticos iniciales reflejen la similitud con el virus HRPLAV (KY052047) y no con el vientovirus LZ posteriormente publicado.

Tras obtener la secuencia completa observamos la presencia de los motivos RCR I,II y III y los motivos Walker A,B y C característicos de la familia *Circoviridae* [238]. En el caso de los redondovirus fueron propuestos otros motivos conservados diferenciados en los géneros Vientovirus y Brisavirus. El estudio de los motivos conservados de estos virus ya indica división entre géneros sin necesidad de hacer filogenia o análisis de distancias.

En el caso de los dos tipos de virus detectados en el estudio de cohorte, el virus HRCiV 15278 presenta motivos conservados más parecidos a los propuestos para el género brisavirus. En el caso del género vientovirus existe más variabilidad en los dominios conservados pero los dominios HuH y SF3 del “Human oral-associated vientovirus LZ” (MK059769) son iguales a los del virus índice HRCiV 15040.

El nonúmero conservado propuesto por este trabajo (ACATTATAT) se encuentra presente en la zona intergénica 1 de todos los virus caracterizados en este trabajo y en todos los virus de la nueva familia *Redondoviridae* [99].

El nonúmero conservado es una de las características que son definitorias para familia dentro de los virus CRESS-DNA, por ello su caracterización es importante a la hora de la descripción de familia en la clasificación taxonómica de los virus CRESS-DNA. El nonúmero propuesto en este

trabajo difiere del propuesto como característico de los redondovirus. Abbas y col. proponen un nonámetro (TATTATTAT) que no está presente en los virus caracterizados en este trabajo y no es común en los virus publicados en su trabajo.

Tras el estudio de virus completo y la caracterización de las estructuras propusimos el nuevo virus “Human Respiratory Circular Virus (HRCiV)”.

Se realizó el estudio de los grupos control para clarificar el papel del virus en el tracto respiratorio. En el caso del grupo 1 solo el 7,6% de las muestras eran positivas. De estas muestras positivas el 17,24% fueron positivas para HRCiV, el resto de los virus presentaban coinfección fundamentalmente con virus ARN. La presencia mayoritaria de coinfección con el virus de la gripe es debido al sesgo del tipo de muestras recibidas en el laboratorio (Tabla 27. Página 156). Debido a su estructura sencilla el estudio de la presencia de HRCiV con otros virus sirve para evaluar si para replicar podría ser un virus dependiente de virus ADN de gran tamaño como los dependoparvovirus [239] o replicaría usando la maquinaria celular como el resto de virus CRESS-DNA [69].

Un trabajo previo ya describió la detección de un nuevo cyclovirus en pacientes pediátricos con infección respiratoria aguda [85]. En este estudio se analizaron 120 aspirados nasofaríngeos y se obtuvieron 4 cyclovirus presentando una prevalencia del 3,3% en el estudio de población. En comparación con el 0,8% de prevalencia de nuestro estudio es una presencia más alta aunque sigue siendo más baja que en adultos confirmando que la presencia de este virus es mayor en adultos que en niños. Este resultado diferencia a estos CRESS-DNA virus de los otros virus circulares como los anelovirus, estos virus se considera que son adquiridos muy poco tiempo después del nacimiento e incluso prenatalmente [232].

Las muestras pediátricas provienen de proyectos acordados con hospitales por lo que se encuentran perfectamente caracterizadas clínicamente y han sido estudiadas y clasificadas desde 1990. En el estudio de la muestras del paciente pediátrico SO 14760 observamos que en la muestra analizada del paciente cuando tenía 3 meses de edad resultó negativa para HRCiV al igual que la de su hermano con 2 años tomada el mismo día que su muestra respiratoria positiva para HRCiV. A pesar de que el número de pacientes analizados son pocos estos resultados pueden dar pistas acerca de cómo se conforma el viroma. Estudios previos establecen que el viroma residente se va conformando en los 2 primeros años de vida como observamos en el estudio de



las muestras de seguimiento por lo que la caracterización de viroma presente en el mismo paciente pediátrico a lo largo del tiempo puede aportar una mejor comprensión de la diversidad viral del tracto respiratorio y la influencia de este viroma en la infección respiratoria [240]. A pesar de su importancia hay que ser cautos a la hora de realizar estudios de viroma respiratorio con pacientes pediátricos ya que la microbiota intestinal no es estable hasta los 3-5 años y puede que no sean extrapolables los resultados obtenidos [241].

La ausencia de muestras positivas en el grupo control sin infección respiratoria tras análisis por PCR HRCiV2 y metagenómica junto con la ausencia de HRCiV en el hermano del paciente pediátrico positivo podría indicar que estos virus no forman parte de la flora comensal. Otros trabajos han caracterizado redondovirus en pacientes sanos [99], quizás la ausencia de nuestros virus en los pacientes control sea debido a que son muestra tomada en pacientes pediátricos por lo que la diferencia de resultados tiene sentido teniendo en cuenta la baja presencia de virus en niños.

La posibilidad de que sea un virus oportunista requiere el estudio de su infección usando otras técnicas como el cultivo celular, el análisis de viremia o la presencia de ARN. Los intentos de cultivo celular resultaron infructuosos con varias líneas celulares empleadas.

El estudio de los virus mostró que dentro de los virus publicados en este trabajo hay dos virus diferentes: HRCiV15040 y HRCiV15278 presentando una homología de secuencia de nucleótidos de genoma completo del 75,6% y por tanto menor del 80 % como indica el ICTV para la consideración de nuevo circovirus.

Como hemos mencionado previamente la publicación de la familia *Redondoviridae* [99] añade un miembro más al grupo de los virus CRESS DNA. Realizando el análisis de distancias y filogenético de los virus HRCiV15040 y HRCiV15278 previamente publicados con los propuestos para esta familia observamos que dos virus son muy parecidos a los caracterizados en esta tesis: “Human oral-associated vintovirus LZ” (MK059769) es similar a HRCiV 15040 y el “Human lung-associated brisavirus II” (MK059755) a HRCiV 15278.

El papel de estos virus en infección respiratoria no está muy claro pero su hallazgo en la mayoría de las muestras de la paciente 1 y que se haya encontrado a niveles más altos en pacientes con infección respiratoria aguda con respecto a muestras de pacientes sanos [99] podría indicar que

al igual que los anelovirus aprovechan periodos de inflamación aguda o inmunosupresión para infectar e incrementar los síntomas [232].

En este objetivo hemos caracterizado virológicamente gracias a la aplicación de la metagenómica clínica en pacientes graves dos virus nuevos que encajan en la nueva familia *Redondoviridae*.

### **5.3 OBJETIVO III: Desarrollo y aplicación de técnicas NGS en el estudio y referencia de los virus gripales y su vigilancia epidemiológica**

En este objetivo se han desarrollado cada uno de los procesos necesarios para implementar de manera exitosa la secuenciación masiva aplicada a un Laboratorio de Referencia de Gripe seleccionando cada una de las técnicas empleadas en el flujo de trabajo para WGS del virus de la gripe.

Las muestras utilizadas en la estandarización de las técnicas WGS de gripe fueron 16 muestras respiratorias (Materiales. Apartado 3.3.4.1) procedentes de la temporada epidémica 2016-2017 y caracterizadas en el laboratorio de Gripe y Virus Respiratorios empleando la PCR tiempo real Gripe A, B y C (Métodos. Apartado 4.12.2).

Con el fin de conocer el impacto del enriquecimiento por 3 métodos en las muestras de gripe se compararon los resultados obtenidos del enriquecimiento viral con 3 métodos y sin enriquecimiento viral en las muestras. En este objetivo fueron empleadas dos muestras positivas a gripe A. Los extractos de RNA de estas muestras fueron sometidos a una amplificación mediante una PCR convencional multisegmento A (Métodos. Apartado 4.7.1) y mediante una PCR tiempo real Gripe A, B y C (Métodos. Apartado 4.12.2).

Para el ajuste de los métodos de amplificación de los segmentos que componen el genoma de gripe se emplearon las diluciones seriadas -2 y -3 de los virus de la gripe de referencia empleados como controles positivos de Gripe A (H1N1pdm), “A/California/07/2009”, y Gripe B (linaje Yamagata), “B/Phuket/3073/2013”, disponibles en el laboratorio. Las diluciones seriadas de estos virus fueron amplificadas empleando una PCR multisegmento Gripe A utilizando la cepa

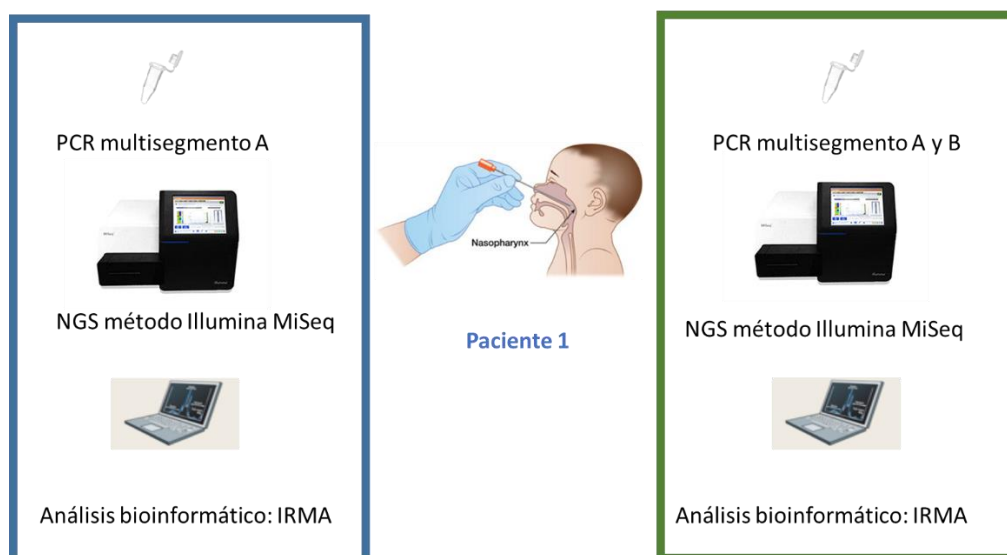
“A/California/07/2009”, una PCR multisegmento Gripe B utilizando la cepa “B/Phuket/3073/2013” y una PCR multisegmento universal para ambos tipos de virus de la gripe.

La disponibilidad de dos de las principales plataformas de secuenciación permitió el análisis comparativo de los resultados WGS de 9 virus de la gripe tipo A/H3N2 previamente caracterizados en el laboratorio mediante la PCR tiempo real Gripe A, B y C (Métodos. Apartado 4.12.2). Tras el empleo del flujo de trabajo para WGS del virus de la gripe (Método. Apartado 4.7) utilizando 2 plataformas con potencial interés para el Laboratorio de Virus Respiratorios y Gripe: 1. el secuenciador *Illumina MiSeq*, disponible en la unidad de Genómica del Instituto de Salud Carlos III, y 2. el secuenciador *MinION* (Materiales. Apartado 3.5.4). El análisis bioinformático de las secuencias obtenidas por WGS se realizó empleando la plataforma de análisis para virus de la gripe *IRMA* descrita en el apartado 4.10.2 de Métodos.

Debido al carácter segmentado del genoma del virus de la gripe el análisis bioinformático de las secuencias obtenidas procedentes de WGS fue realizado en plataformas específicas para este virus. Las plataformas empleadas fueron: *INSaFLU* [25] e *IRMA* [21] (Métodos. Apartado 4.10.2). Se analizaron las secuencias procedentes de 6 muestras a las que les fueron detectadas Gripe A/H3N2 y Gripe B, a estas muestras se les realizó el flujo de trabajo para WGS del virus de la gripe (Métodos. Apartado 4.7).

Procedentes del Centro Nacional de Gripe de Valladolid fueron recibidas 12 muestras (Tabla 8. Página 73) de una paciente pediátrica a la que se le detectó la presencia de Gripe B perteneciente al linaje Yamagata. El historial clínico de la paciente refirió fiebre, catarro, otalgia que evolucionó a un cuadro complejo que presentaba signos neurológicos e hipotensión. Una semana después de la primera visita al Centro de Salud la paciente sufrió un síncope que terminó en *exitus* de la paciente (Materiales. Apartado 3.3.4.2).

Tras la estandarización de los métodos necesarios para la secuenciación WGS del virus de la gripe, se realizó un estudio de las 12 muestras de la paciente pediátrica (Figura 28).



**Figura 28:** Flujo de trabajo para WGS del virus de gripe empleado en las muestras de la paciente 1

El extracto de las muestras de la paciente pediátrica 1 fue analizado con la PCR multisegmento A (Métodos. Apartado 4.7.1). Con el producto de amplificación obtenido se prepararon las librerías empleando el kit *Nextera XT* (Métodos. Apartado 3.5.2) para su posterior secuenciación con la plataforma *MiSeq* (Métodos. Apartado 3.5.4).

El hermano de la paciente 1, el paciente pediátrico 2, presentó también un cuadro respiratorio con otalgia y fiebre. Dada la gravedad de la infección de la paciente 1, fueron enviadas desde el Centro Nacional de Gripe de Valladolid, 1 muestra respiratoria y 2 aislamientos virales en células MDCK del virus del paciente 2. A este paciente le fue detectada una coinfección de gripe A y B por métodos de diagnóstico molecular.

Las muestras de la paciente 1 y del paciente 2 fueron de nuevo analizadas empleando esta vez las PCRs multisegmento A y B (Métodos. Apartado 4.7.1 y 4.7.2.). La preparación de librerías se realizó con el kit *Nextera XT* (Métodos. Apartado 3.5.2) y fueron secuenciadas con el secuenciador *MiSeq* (Métodos. Apartado 3.5.4) (Figura 28. Página 168).

En este objetivo se desarrolló el flujo de trabajo para WGS del virus de la gripe que permitió el estudio de un caso clínico de una paciente pediátrica y su hermano que presentaban una coinfección de gripe B y gripe A/H3N2.

### 5.3.1. Desarrollo de protocolos para el uso de WGS en el laboratorio de Referencia de Gripe

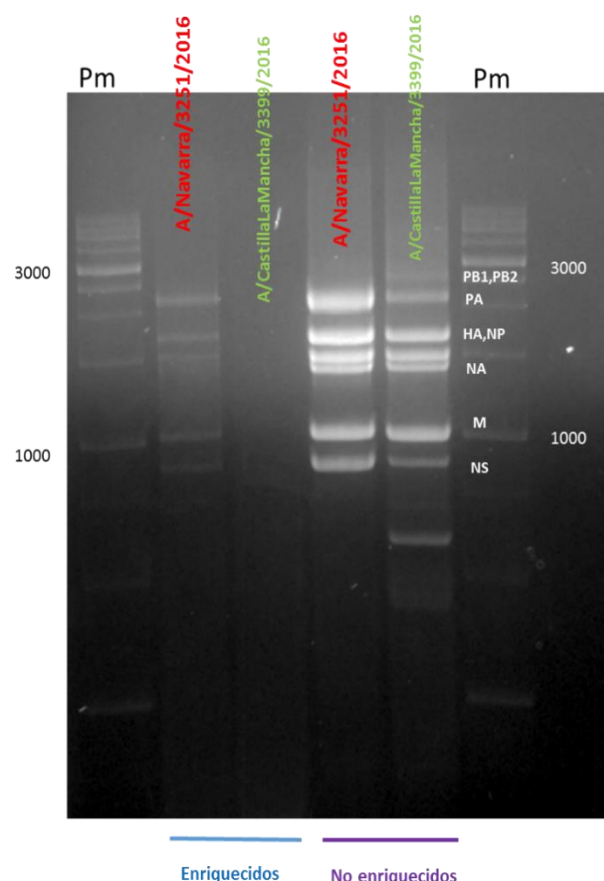
En este apartado centrado en la estandarización de protocolos para la secuenciación WGS del virus de la gripe se emplearon 2 muestras respiratorias positivas para virus de la gripe A (H3N2), “A/Navarra/3251/2016” y “A/CastillaLaMancha/3399/2016”, (Tabla 7. Página 72) para la evaluación del impacto del protocolo de enriquecimiento con 3 métodos (Métodos. Apartado 4.5.1) en la secuenciación WGS del virus de la gripe.

El extracto de las muestras duplicadas, tras el enriquecimiento viral y sin enriquecimiento viral, fue sometido a dos protocolos de PCR convencional: PCR tiempo real de gripe A, B y C (Métodos. Apartado 4.12.2) con el fin de valorar la presencia de virus en base a su umbral de ciclo (ct) y la PCR multisegmento A que tiene como objetivo la secuenciación del genoma completo del virus de la gripe A [16].

El análisis realizado mediante electroforesis en un gel de agarosa al 1% de los productos de amplificación de la PCR multisegmento A reflejó que la muestra enriquecida presentaba menor cantidad de ADN amplificado cuando se comparaba con la muestra no enriquecida. Este resultado se reflejó en ambas muestras analizadas (Figura 28. Página 168).

En el análisis de la cuantificación por PCR en tiempo real se obtuvieron resultados similares al experimento anterior. El virus de la gripe “A/Navarra/3251/2016” presentó una diferencia de 3 ciclos entre la muestra no enriquecida y la enriquecida y en el virus de la gripe “A/CastillaLaMancha/3399/2016” la diferencia entre la muestra no enriquecida y la enriquecida fue de 7 ciclos (Tabla 31).

Teniendo en cuenta ambos resultados se concluyó que para obtener una mayor sensibilidad del procedimiento no es necesario el enriquecimiento viral de las muestras para la WGS del virus de la gripe.



**Figura 29:** Resultados comparativos entre muestras enriquecidas y no enriquecidas. Gel de agarosa al 1% comparando los virus de Gripe A (H3N2) “A/Navarra/3251/2016” y “A/CastillaLaMancha/3399/2016” tras realizarse el protocolo de enriquecimiento con 3 métodos y sin la realización de éste.

Identificación del virus	Ct muestras enriquecidas viralmente	Ct de ciclo muestras no enriquecidas viralmente
A/Navarra/3251/2016	24,6	27,5
A/CastillaLaMancha/3399/2016	23,2	30,4

**Tabla 31:** Umbrales de ciclo de los virus de la gripe analizados obtenidos tras la PCR tiempo real Gripe A, B y C en muestras enriquecidas viralmente y sin enriquecimiento viral

### 5.3.2 Amplificación de los ácidos nucleicos: selección de PCR multisegmento

Tres métodos han sido previamente publicados para la amplificación simultánea de los ocho segmentos del genoma del virus de la gripe. Los métodos descritos son la PCR multisegmento para Gripe A [16], la PCR multisegmento para Gripe B [17] y la PCR multisegmento universal para la amplificación de los 8 segmentos de los virus de la gripe tipo A y B simultáneamente [18].

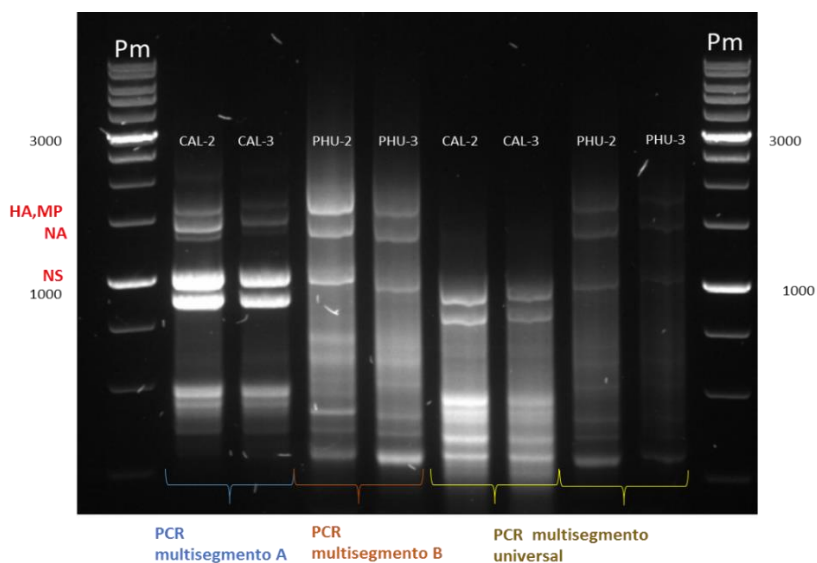
Con el fin de estandarizar estas 3 PCR convencionales en el laboratorio de Gripe y Virus Respiratorios fueron empleadas las diluciones -2 y -3 de los virus empleados como controles positivos de Gripe A (H1N1pdm), “A/California/07/2009”, y Gripe B (linaje Yamagata), “B/Phuket/3073/2013”, disponibles en el laboratorio.

La electroforesis en un gel de agarosa al 1% de las reacciones de amplificación de las 3 PCR empleadas, mostró que las PCR multisegmento por tipo de gripe (multisegmento A y B) son métodos más sensibles para la secuenciación de los segmentos de gripe que la PCR multisegmento universal. Especialmente en el caso del virus de la gripe B (Figura 30), la PCR multisegmento B mostró mejor resultado en la amplificación de los segmentos del virus que el método de amplificación universal para ambos virus de la gripe.

### 5.3.3 Selección de la plataforma de secuenciación: *Illumina MiSeq* vs *MinION*

En este objetivo se secuenciaron 10 virus de la gripe detectados en muestras procedentes de la Red de Vigilancia de Gripe de la temporada epidémica 2016-2017 (Tabla 8. Página 73) empleando las plataformas *Illumina MiSeq* y *MinION*.

El análisis bioinformático de los resultados obtenidos mostró que las secuencias consenso procedentes de la secuenciación por la plataforma *MiSeq* y las secuencias consenso procedentes de la secuenciación en la plataforma *MinION* coincidían en un 99%-100% tanto en tamaño de secuencia como en homología en todos los segmentos (Tabla 33).



**Figura 30.** Gel de agarosa al 1% que muestra los resultados obtenidos en la amplificación por PCR multisegmento A (llave azul), B (llave naranja) y la PCR multisegmento universal (llaves amarillas) usando la dilución -2 y -3 de los virus A (H1N1 pdm) y B (linaje Yamagata). CAL: “A/California/07/2009”. PHU: B/Phuket/3073/2013

Proteína Básica 2 de la Polimerasa PB2		
Muestras	Tamaño de secuencia Minlon/Illumina	Homología (%)
A/Navarra/3251/2016	2280 / 2280	99,94
A/Cantabria/726/2017	2279 / 2280	100
A/Extremadura/390/2017	2279 / 2280	99,94
A/Ceuta/170026/2016	2278 / 2280	100
A/Cantabria/1387/2017	2276 / 2280	99,82
A/Extremadura/1026/2017	2280 / 2280	100
A/CastillaLaMancha/769/2017	2280 / 2280	100
A/Cantabria/730/2017	2280 / 2280	99,94
A/LaRioja/170884/2016	2308/2311	98,91%
Proteína Básica 1 de la Polimerasa PB1		
Muestras	Tamaño de secuencia Minlon/Illumina	Homología (%)
A/Navarra/3251/2016	N/A	N/A
A/Cantabria/726/2017	2272 / 2274	99,90
A/Extremadura/390/2017	2274 / 2274	100
A/Ceuta/170026/2016	N/A	N/A
A/Cantabria/1387/2017	2316/1484	89,86
A/Extremadura/1026/2017	2271 / 2274	99,87
A/CastillaLaMancha/769/2017	2274 / 2274	100
A/Cantabria/730/2017	2273 / 2274	99,96
A/LaRioja/170884/2016	2270 / 2274	99,82
Polimerasa Ácida PA		
Muestras	Tamaño de secuencia Minlon/Illumina	Homología (%)
A/Navarra/3251/2016	2150 / 2151	99,95
A/Cantabria/726/2017	2151 / 2151	100
A/Extremadura/390/2017	2151 / 2151	100
A/Ceuta/170026/2016	2150 / 2151	99,95
A/Cantabria/1387/2017	2151 / 2151	100
A/Extremadura/1026/2017	2150 / 2151	99,95
A/CastillaLaMancha/769/2017	2151 / 2151	100
A/Cantabria/730/2017	2150 / 2151	99,95
A/LaRioja/170884/2016	2150 / 2151	99,95
Hemaglutinina HA		
Muestras	Tamaño de secuencia Minlon/Illumina	Homología (%)
A/Navarra/3251/2016	1700 / 1701	99,94
A/Cantabria/726/2017	1701 / 1701	100
A/Extremadura/390/2017	1700 / 1701	99,94
A/Ceuta/170026/2016	1701 / 1701	100
A/Cantabria/1387/2017	1698 / 1701	99,82
A/Extremadura/1026/2017	1701 / 1701	100
A/CastillaLaMancha/769/2017	1701 / 1701	100
A/Cantabria/730/2017	1700 / 1701	99,94
A/LaRioja/170884/2016	1701 / 1701	100

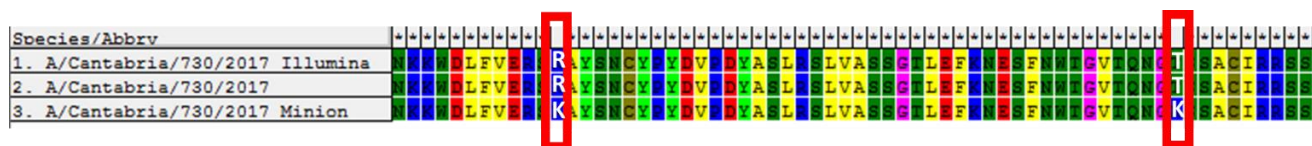


Nucleoproteína NP		
Muestras	Tamaño de secuencia Minlon/Illumina	Homología (%)
A/Navarra/3251/2016	1497 / 1497	100
A/Cantabria/726/2017	1497 / 1497	100
A/Extremadura/390/2017	1496 / 1497	99,93
A/Ceuta/170026/2016	1494 / 1497	99,80
A/Cantabria/1387/2017	N/A	N/A
A/Extremadura/1026/2017	1497 / 1497	100
A/CastillaLaMancha/769/2017	1496 / 1497	99,93
A/Cantabria/730/2017	1497 / 1497	100
A/LaRioja/170884/2016	1497 / 1497	100
Neuraminidasa NA		
Muestras	Tamaño de secuencia Minlon/Illumina	Homología (%)
A/Navarra/3251/2016	1410 / 1410	100
A/Cantabria/726/2017	1410 / 1410	100
A/Extremadura/390/2017	1410 / 1410	100
A/Ceuta/170026/2016	1410 / 1410	100
A/Cantabria/1387/2017	1406 / 1410	99,65
A/Extremadura/1026/2017	1410 / 1410	100
A/CastillaLaMancha/769/2017	1410 / 1410	100
A/Cantabria/730/2017	1410 / 1410	100
A/LaRioja/170884/2016	1408 / 1410	99,86
Matriz M		
Muestras	Tamaño de secuencia Minlon/Illumina	Homología (%)
A/Navarra/3251/2016	982 / 982	100
A/Cantabria/726/2017	982 / 982	100
A/Extremadura/390/2017	981 / 982	99,90
A/Ceuta/170026/2016	981/982	99,90
A/Cantabria/1387/2017	980 / 982	99,8
A/Extremadura/1026/2017	981 / 982	99,90
A/CastillaLaMancha/769/2017	982 / 982	100
A/Cantabria/730/2017	981 / 982	99,90
A/LaRioja/170884/2016	982 / 982	100
Proteína no estructural NS		
Muestras	Tamaño de secuencia Minlon/Illumina	Homología (%)
A/Navarra/3251/2016	837 / 838	99,88
A/Cantabria/726/2017	838 / 838	100
A/Extremadura/390/2017	837 / 838	99,88
A/Ceuta/170026/2016	838 / 838	100
A/Cantabria/1387/2017	836 / 838	99,76
A/Extremadura/1026/2017	837 / 838	99,88
A/CastillaLaMancha/769/2017	838 / 838	100
A/Cantabria/730/2017	838 / 838	100
A/LaRioja/170884/2016	838 / 838	100

**Tabla 32:** Resultados obtenidos del tamaño y la homología de las secuencias consenso obtenido, empleando los secuenciadores *MiniON* y *MiSeq Illumina*

A pesar de la gran similitud de la secuencias consenso de ambas plataformas, se detectaron en todos los segmentos de los virussecuenciados en la plataforma *MinION* variantes minoritarias de nucleótidos soportadas por decenas o centenares de lecturas que no estaban presentes en las secuencias consenso obtenidas con el secuenciador *MiSeq*.

Debido a la importancia que pueden tener la aparición de estos cambios de nt en las secuencias consenso, especialmente en el caso de HA, fueron comparadas las secuencias obtenidas por WGS con las obtenidas por el método de referencia que es el método Sanger (Métodos. Apartado 4.14). Las variantes encontradas en las secuencias obtenidas usando el secuenciador *MinION* no se encontraron en la secuencia obtenida por el método Sanger. Estos cambios de nt erróneos en las secuencias obtenidas de la plataforma *MinION* en algunos casos causaban cambio de aminoácidos (Figura 14. Página 117). Por el contrario, las secuencias consenso obtenidas usando el secuenciador *MiSeq* y los fragmentos obtenidos por Sanger fueron coincidentes al 100% mostrando que el empleo del método *Illumina* con la plataforma *MiSeq* debería ser el método de elección en el laboratorio de salud pública (Figura 31).



**Figura 31:** Fragmento del alineamiento de las secuencias de aminoácidos de las secuencias consenso de la HA de gripe obtenidas en la secuenciación con *MiSeq*, *MinION* y el método Sanger. El alineamiento fue realizado con el programa *Clustal X2* y la traducción a aminoácidos del alineamiento fue realizado con el software *MEGA 5.2*

### 5.3.4 Análisis bioinformático de la WGS del virus de la gripe

Se analizaron en ambas plataformas las secuencias obtenidas del análisis de 3 virus gripales A/H3N2, “A/Galicia/2444/2017”, “A/Galicia/2467/2017”, “A/Baleares/2479/2017” y 3 virus gripales tipo B “B/Galicia/2406/2017”, “B/Galicia/2407/201” y “B/CastillaLaMancha/2439/2017” previamente caracterizados empleando la PCR tiempo real Gripe A, B y C (Métodos. Apartado 4.12.2).

Uno de los requisitos de la plataforma *INSaFLU* es la necesidad de introducir por parte del usuario un virus de referencia para la generación de secuencias consenso de las muestras a analizar (Métodos. 4.10.2).


Debido a que los virus empleados en este análisis fueron previamente caracterizados en el Laboratorio de Gripe y Virus Respiratorios para el análisis de la temporada epidémica de gripe 2016/2017, se disponía de la secuencia de un fragmento de HA de los virus obtenida por secuenciación por Sanger. Se realizó un análisis empleando la herramienta *BLASTn* (Materiales. Apartado 3.5.5) del fragmento de la HA de los virus de interés para la obtención de virus con la mayor homología posible que pudiesen ser empleados como virus de referencia para el ensamblaje de las secuencias (Tabla 33).

Durante el análisis realizado empleando la plataforma *INSaFLU* fueron obtenidas las secuencias consenso de los 8 segmentos de los 3 virus de la gripe A/H3N2 secuenciados (Figura 32). En el caso de los virus de la gripe B no pudieron ser obtenidos inicialmente las secuencias consenso de los 8 segmentos de gripe de los 3 virus de interés.

Identificación del virus	Virus de referencia en <i>INSaFLU</i>	Segmentos del genoma obtenidos
<b>A/Galicia/2444/2017</b>	A/HongKong/4801/2014	Genoma completo
<b>A/Galicia/2467/2017</b>	A/HongKong/4801/2014	Genoma completo
<b>A/Baleares/2479/2017</b>	A/HongKong/4801/2014	Genoma completo
<b>B/Galicia/2406/2017</b>	B/Norway/2409/2017	Genoma completo
<b>B/Galicia/2407/201</b>	B/Norway/2409/2017	Genoma completo
<b>B/CastillaLaMancha/2439/2017</b>	B/Norway/2409/2017	PB2, PB1, HA, NP, NA, M, NS
<b>B/Galicia/2406/2017</b>	B/Brisbane/60/2008	PB2, PB1, PA, NP, NA, M, NS
<b>B/Galicia/2407/201</b>	B/Brisbane/60/2008	PB2, PB1, PA, NP, NA, M, NS
<b>B/CastillaLaMancha/2439/2017</b>	B/Brisbane/60/2008	PB2, PB1, PA, NP, NA, M, NS

**Tabla 33:** Segmentos del genoma obtenido de los diferentes virus analizados con las secuencias de referencia para el ensamblaje empleadas en *INSaFLU*

En el caso de los virus de la gripe B a analizar el virus con más homología disponible en la base de datos *GenBank* (Materiales. Apartado 3.5.6) fue el virus “B/Norway/2409/2017” por lo que fue empleado como virus de referencia para el ensamblaje en la plataforma *INSaFLU*.

Sample Name	Type And Subtype	Putative Mixed Infection	Dataset	Coverage	Consensus File
 172444	A-N2	No	Generic		<a href="#">Consens...44.fasta</a>
 172467	A-N2	No	Generic		<a href="#">Consens...67.fasta</a>
 172479	A-N2	No	Generic		<a href="#">Consens...79.fasta</a>

**Figura 32:** Gráfico generado por *INSaFLU* que muestra la cobertura y secuenciación de los ocho segmentos de gripe de los 3 virus de gripe A analizados. 172444: “A/Galicia/2444/2017”. 172467: “A/Galicia/2467/2017”. 172479: “A/Baleares/2479/2017”. Cada punto verde representa la generación exitosa de secuencia consenso para ese segmento del genoma del virus.

En el análisis de los virus “B/Galicia/2406/2017” y “B/Galicia/2407/2017” empleando como referencia el virus “B/Norway/2409/2017” se obtuvieron las secuencias consenso completas de los 8 segmentos de los virus “B/Galicia/2406/2017” y “B/Galicia/2407/2017” (Figura 33).

Sample Name	Type And Subtype	Putative Mixed Infection	Dataset	Coverage	Consensus File	Alerts
172407	B-Victoria	No	Generic	●●●●●●●●	<a href="#">Consens...07.fasta</a>	0
172406	B	No	Generic	●●●●●●●●	<a href="#">Consens...06.fasta</a>	0
172439	B-Victoria	No	Generic	●●●●●●●●	<a href="#">Consens...39.fasta</a>	2

**Figura 33:** Gráfico generado por *INSaFLU* que muestra la cobertura y secuenciación de los ocho segmentos de gripe de los 3 virus de gripe B analizados empleando como secuencia de referencia “B/Norway/2409/2017”. 172407: “B/Galicia/2407/2017”. 172406: “B/Galicia/2406/2017”. 172439: “B/CastillaLaMancha/2439/2017”. Los puntos naranjas representan la generación de la secuencia consenso incompleta del segmento que representa.

Para contrastar la importancia de la elección de la secuencia de referencia para el ensamblaje en la plataforma *INSaFLU* fueron analizados los virus “B/Galicia/2406/2017” y “B/Galicia/2407/2017” usando como referencia la cepa vacunal del linaje Victoria de la temporada 2016/2017: el virus “B/Brisbane/60/2008” [108]. El empleo de la secuencia del virus “B/Brisbane/60/2008” aportó peores resultados, no obteniéndose la secuencia consenso del segmento HA en los virus “B/Galicia/2406/2017” y “B/Galicia/2407/2017” que previamente había podido ser ensamblada empleando la secuencia del virus “B/Norway/2409/2017” como referencia (Figura 34).

Sample Name	Type And Subtype	Putative Mixed Infection	Dataset	Coverage	Consensus File	Alerts
172406	B	No	Generic	●●●●●●●●	<a href="#">Consens...06.fasta</a>	1
172407	B-Victoria	No	Generic	●●●●●●●●	<a href="#">Consens...07.fasta</a>	1
172439	B-Victoria	No	Generic	●●●●●●●●	<a href="#">Consens...39.fasta</a>	2

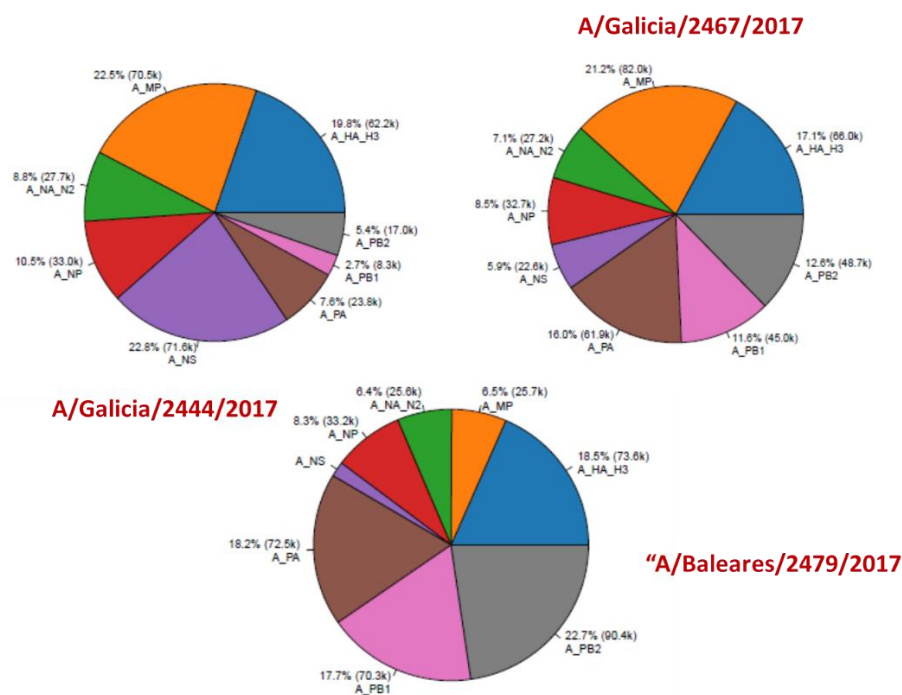
**Figura 34:** Gráfico generado por *INSaFLU* que muestra la cobertura y secuenciación de los ocho segmentos de gripe de los 3 virus de gripe B analizados empleando como secuencia de referencia “B/Brisbane/60/2008”. 172407: “B/Galicia/2407/2017”. 172406: “B/Galicia/2406/2017”. 172439: “B/CastillaLaMancha/2439/2017”. El punto rojo representa la no generación de secuencia consenso para el segmento del genoma de gripe que representa.

El análisis del virus “B/CastillaLaMancha/2439/2017” arrojó resultados diferentes según el virus de referencia introducido por el usuario en la plataforma para el ensamblaje de los fragmentos obtenido por WGS. La secuencia consenso del segmento PB2 generada no cubría el gen completo en ninguna de las secuencias de referencia empleadas para el ensamblaje. Tras el análisis empleando como secuencia de referencia al virus “B/Norway/2409/2017” no se generó la secuencia consenso completa del segmento PA, sin embargo tras el análisis de las secuencias del virus “B/CastillaLaMancha/2439/2017” empleando como referencia la secuencia de la cepa vacunal “B/Brisbane/60/2008” se generó, a diferencia de lo ocurrido con los virus “B/Galicia/2406/2017” y “B/Galicia/2407/2017”, la secuencia consenso para HA aunque no se obtuvo la secuencia del gen completo.

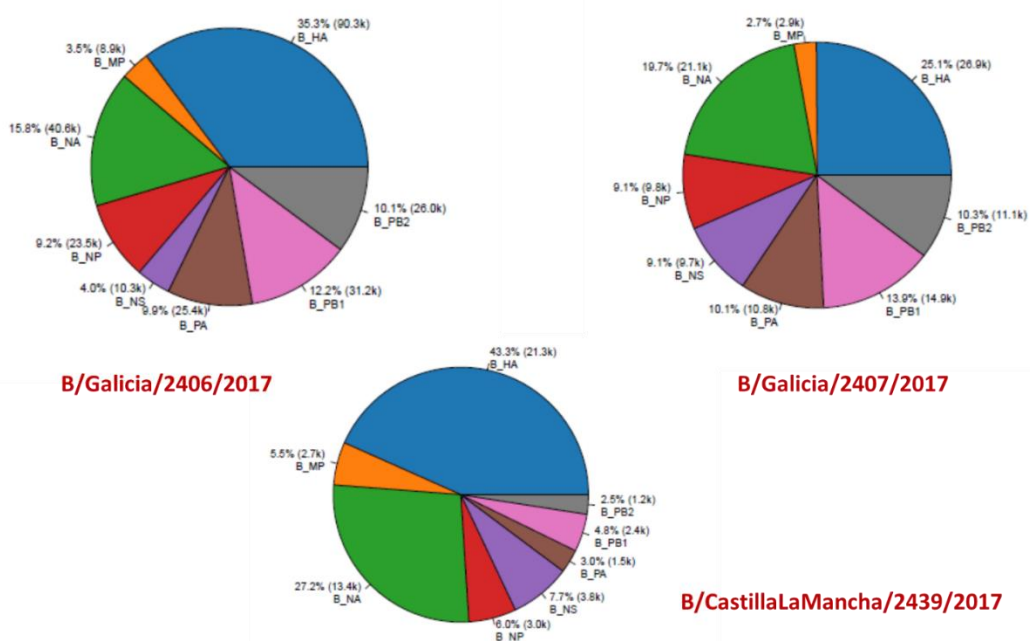
Una de las grandes diferencias entre las plataformas empleadas es que el método de ensamblaje utilizado por la plataforma IRMA no requiere la introducción de una secuencia de referencia por el usuario.

Los resultados obtenidos para los 3 virus gripales A/H3N2 analizados, “A/Galicia/2444/2017”, “A/Galicia/2467/2017”, “A/Baleares/2479/2017”, mostraron la generación de las secuencias del genoma completo de estos virus. Las secuencias consenso obtenidas por *IRMA* y por *INSaFLU* de estos virus fueron idénticas (Figura 35).

En el caso de las 3 gripes B analizadas, “B/Galicia/2406/2017”, “B/Galicia/2407/2017” y “B/CastillaLaMancha/2439/2017”, también fueron obtenidos los 8 segmentos del genoma de gripe completos (Figura 36).



**Figura 35:** Gráfico generado por IRMA de la distribución de lecturas de los 8 segmentos de los virus de la gripe A “A/Galicia/2444/2017”, “A/Galicia/2467/2017”, “A/Baleares/2479/2017”



**Figura 36:** Gráfico generado por IRMA de la distribución de lecturas de los 8 segmentos de los virus de la gripe B, “B/Galicia/2406/2017”, “B/Galicia/2407/2017” y “B/CastillaLaMancha/2439/2017”.

Todas las secuencias de virus de la gripe obtenidas durante la estandarización de los métodos WGS para gripe fueron depositadas en la base de datos *GISAID* (Tabla 34) (Materiales. Apartado 3.5.6)

Debido a los resultados previamente descritos la elección de la plataforma *IRMA* se ajusta más a las necesidades del laboratorio de Virus Respiratorios al proporcionarnos las secuencias completas de manera fiable, sin necesidad de conocer previamente la referencia y aportando además de la secuencia consenso información sobre la diversidad viral en la muestra secuenciada

Cepa de virus de la gripe	Tipo y subtipo	Identificación de GISAID
A/Navarra/3251/2016	A/H3N2	EPI_ISL_242989
A/CastillaLaMancha/3399/2016	A/H3N2	EPI_ISL_242976
A/Navarra/3251/2016	A/H3N2	EPI_ISL_242989
A/Cantabria/726/2017	A/H3N2	EPI_ISL_252954
A/Extremadura/390/2017	A/H3N2	EPI_ISL_258083
A/Ceuta/170026/2016	A/H3N2	EPI_ISL_242978
A/Cantabria/1387/2017	A/H3N2	EPI_ISL_258428
A/Extremadura/1026/2017	A/H3N2	EPI_ISL_256436
A/CastillaLaMancha/769/2017	A/H3N2	EPI_ISL_253432
A/Cantabria/730/2017	A/H3N2	EPI_ISL_252958
A/LaRioja/170884/2016	A/H3N2	EPI_ISL_256437
A/Galicia/2444/2017	A/H3N2	EPI_ISL_288740
A/Galicia/2467/2017	A/H3N2	EPI_ISL_291781
A/Baleares/2479/2017	A/H3N2	EPI_ISL_291772
B/Galicia/2406/2017	B/Victoria	EPI_ISL_288426
B/Galicia/2407/2017	B/Victoria	EPI_ISL_288427
B/CastillaLaMancha/2439/2017	B/Victoria	EPI_ISL_288741

**Tabla 34:** Secuencias de virus obtenidas por WGS y número de identificación de los virus depositados en la base de datos *GISAID*

### 5.3.5 Análisis del caso clínico de paciente pediátrica con gripe

Para el análisis de las 12 muestras de la paciente pediátrica 1, todas las PCRs múltiples tiempo real (Métodos. Apartado 4.12) disponibles en el Laboratorio de Virus Respiratorios y Gripe fueron realizadas para completar el diagnóstico de manera homogénea. Los resultados obtenidos del



análisis de las muestras fueron negativos para las PCRs múltiples BRQ y Respiratorios 2 y en la muestra correspondiente al hisopo nasofaríngeo fue detectado un virus de la gripe tipo B (Tabla 35).

Identificación muestras	PCR Gripe A, B y C	PCRs BRQ/Respiratorios 2	PCR Multisegmento B	PCR Multisegmento A
181425	Negativo	Negativo	Negativo	Negativo
181426	Negativo	Negativo	Negativo	Negativo
181427	Negativo	Negativo	Negativo	Negativo
181428	Negativo	Negativo	Positivo	Negativo
181429	Negativo	Negativo	Negativo	Negativo
181430	Negativo	Negativo	Positivo	Positivo
181431	Negativo	Negativo	Negativo	Negativo
181432	Negativo	Negativo	Negativo	Negativo
181433	Negativo	Negativo	Positivo	Positivo
181434	Positivo (B)	Negativo	Positivo	Positivo
181435	Negativo	Negativo	Positivo	Positivo
181436	Negativo	Negativo	Positivo	Positivo

**Tabla 35:** Resultados del análisis de las 12 muestras procedentes de la paciente pediátrica tras realizarles las PCRs tiempo real Gripe A, B y C, múltiple BRQ, múltiple Respiratorios 2 y las PCRs convencionales multisegmento A y B.

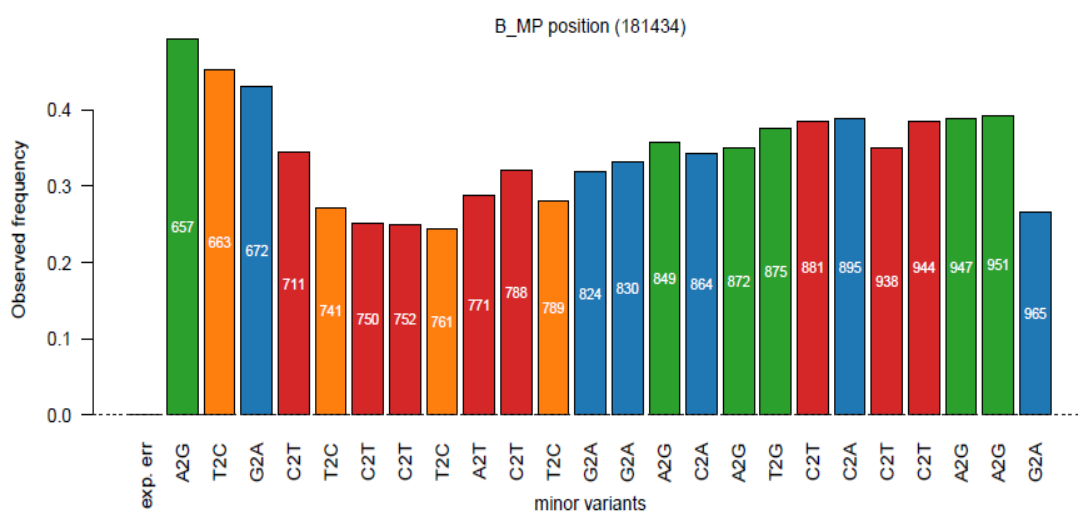
Debido a la gravedad del caso se decidió realizar un análisis por WGS del virus detectado en la paciente con el fin de caracterizar todos los segmentos del virus. Siguiendo el flujo de trabajo para WGS de gripe (Métodos. Apartado 4.7) se realizó la PCR multisegmento B (Métodos. Apartado 4.7.2) en los extractos de las muestras de la paciente detectándose virus de la gripe en 6 de las 12 muestras procedentes de la paciente (2 lóbulos del pulmón, hisopo faríngeo, hisopo nasofaríngeo, oído derecho y oído izquierdo) (Tabla 8. Página 73).

Los productos de amplificación de la PCR multisegmento B fueron secuenciados empleando la preparación de librerías con el kit *Nextera XT* (Materiales. Apartado 3.5.2) y utilizando la plataforma de secuenciación *MiSeq* (Materiales. Apartado 3.5.4).

El análisis bioinformático de las secuencias se realizó empleando la plataforma *IRMA* (Materiales. Apartado 3.5.5) que mostró que 6 de las 12 muestras eran positivas al virus de la gripe B (Tabla 22. Página 135). Estos virus fueron caracterizados como gripe B y asociado genéticamente a la cepa “B/Phuket/3073/2013” referencia del linaje Yamagata. La secuencia de este virus fue depositada en la base de datos GISAID (Materiales. Apartado 3.5.6), “B/Palencia/1434/2018” (EPI\_ISL\_303626).



Analizando cada uno de los segmentos del virus “B/Palencia/1434/2018” se observó en el segmento M numerosos cambios con respecto la secuencia consenso en una proporción entre el 30%-50% (Figura 37). Con el fin de caracterizar el origen de los cambios en este segmento se realizaron 3 aproximaciones: 1. análisis de las secuencias secundarias detectadas por IRMA, 2. segundo análisis bioinformático empleando otro método de ensamblaje alternativo, y 3. Repetición de la secuenciación WGS de la muestra 181434 para descartar posibles contaminaciones.



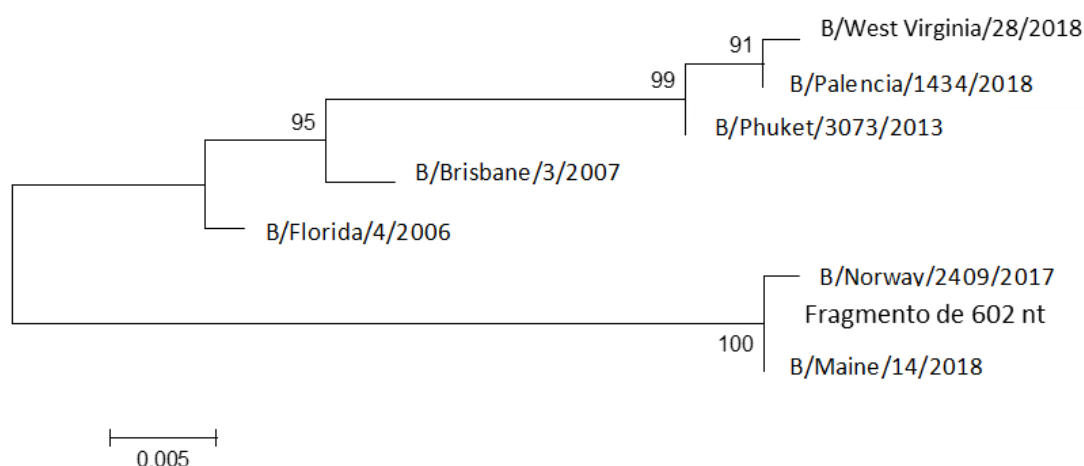
**Figura 37:** Diagrama de cambio de nucleótidos y las frecuencias en la que se producen en el segmento M del virus “B/Palencia/1434/2018”. Gráfico generado por la plataforma IRMA

En el análisis de las secuencias secundarias se observó la presencia de un fragmento de 302 nt que fue identificado como gen M del virus de la gripe A, así como un fragmento de 274 nt del gen HA del virus de la gripe A/H3. Debido a este hallazgo se realizó un alineamiento empleando el programa de alineamiento de secuencias *MACAW* entre los fragmentos detectados de M de gripe A y varios virus gripe A/H3N2 circulantes en la temporada epidémica 2016/2017. En este análisis se observó un alineamiento perfecto entre los fragmentos de 302 nt del gen M de gripe A y el gen M del virus de referencia A/H3N2 “A/Singapore/INFIMH-16-0019/2016” (Figura 38). La presencia de estos fragmentos requería un análisis posterior para descartar una posible contaminación de la muestra.



**Figura 38:** Alineamiento realizado con el programa *MACAW* entre los fragmentos de M gripe A detectados en la muestras 181434 y la gripe A H3N2 “A/Singapore/INFIMH-16-0019/2016”

La segunda aproximación consistió en contrastar los resultados con los obtenidos en un análisis basado en el ensamblaje por el método de *síntesis de novo* realizado en la plataforma *Virus TAP* [24]. El análisis por *síntesis de novo* detectó un fragmento de 602 nt con una profundidad de lectura de 19x del gen M del virus de la Gripe B filogenéticamente relacionado con el linaje Victoria en vez de con el linaje Phuket al que pertenecía el virus consenso (Figura 39).



**Figura 39:** Análisis filogenético realizado por el modelo Hasegawa-Kishino-Yano del método *Maximum Likelihood* con un respaldo estadístico de 5000 repeticiones. En el que se muestra la agrupación del fragmento de matriz obtenido con el grupo de la gripe circulante “B/Norway/2409/2017” del linaje Victoria en vez de con el grupo de “B/Phuket/3073/2013” del linaje Yamagata con el que se agrupaba en el resto de segmentos.

Debido la obtención de resultados contradictorios como la presencia de un fragmento de matriz de Gripe A o un fragmento de matriz de Gripe B linaje Victoria se volvió a realizar el flujo de trabajo WGS empleando la PCR multisegmento B en las 12 muestras de la paciente 1. Tras repetir la secuenciación de esta muestra no se volvió a obtener el fragmento de M de Gripe B del linaje Victoria pero se obtuvo de nuevo un fragmento de 292 nt de M de gripe A/H3N2 (Figura 40).



**Figura 40:** Alineamiento realizado con el programa *MACAW* entre los fragmentos de M gripe A detectados en la muestra 181434 tras los dos métodos WGS realizados y la gripe A H3N2 “A/Singapore/INFIMH-16-0019/2016”

Durante el estudio de la presencia de una posible coinfección de Gripe B linaje Yamagata y Gripe A/H3N2 en las muestras respiratorias de la paciente pediátrica 1 se recibió procedente del Centro Nacional de Gripe de Valladolid una muestra de ENF del hermano de la paciente en el que se había detectado una coinfección de gripe B y gripe A/H3N2 confirmada en dicho laboratorio (Tabla 9. Página 73).

Identificación de la muestra	PCR Gripe A, B y C	PCRs BRQ Respiratorios 2	PCR Multisegmento B	PCR Multisegmento A
<b>181434B</b>	<b>Positivo (A y B)</b>	N/A	<b>Positivo</b>	<b>Positivo</b>
<b>181434B-MDCK-SIAT</b>	<b>Positivo (A y B)</b>	N/A	<b>Positivo</b>	<b>Positivo</b>
<b>181434B-MDCK</b>	<b>Positivo (A y B)</b>	N/A	<b>Positivo</b>	<b>Positivo</b>

**Tabla 36:** Resultados del análisis de las muestras procedentes del paciente 2 tras análisis mediante las PCRs tiempo real Gripe A, B y C, múltiple BRQ, múltiple Respiratorios 2 y las PCRs convencionales multisegmento A y B.

Para completar la caracterización virológica de las muestras del paciente 2 fueron realizadas las PCRs tiempo real de diagnóstico del laboratorio (Métodos. Apartado 4.12) resultando negativas para las PCRs múltiples tiempo real BRQ y Respiratorios 2. En este caso fueron detectados virus de la gripe tipo A y B por la PCR tiempo real Gripe A, B y C (Tabla 36) (Métodos. Apartado 4.12.2).

Con el fin de caracterizar los virus de ambos pacientes y de confirmar la coinfección en el caso de la paciente 1 se realizaron las PCRs Multisegmento B [17] y Multisegmento A [16] dentro del flujo de trabajo para WGS de gripe (Métodos. Apartado 4.7).

El análisis bioinformático de la WGS del virus de la gripe en el caso de la paciente 1 demostró que los resultados obtenidos empleando la PCR multisegmento B eran los mismos, confirmándose la presencia de un fragmento del gen M de gripe A. Tras el análisis por la plataforma *IRMA* de las

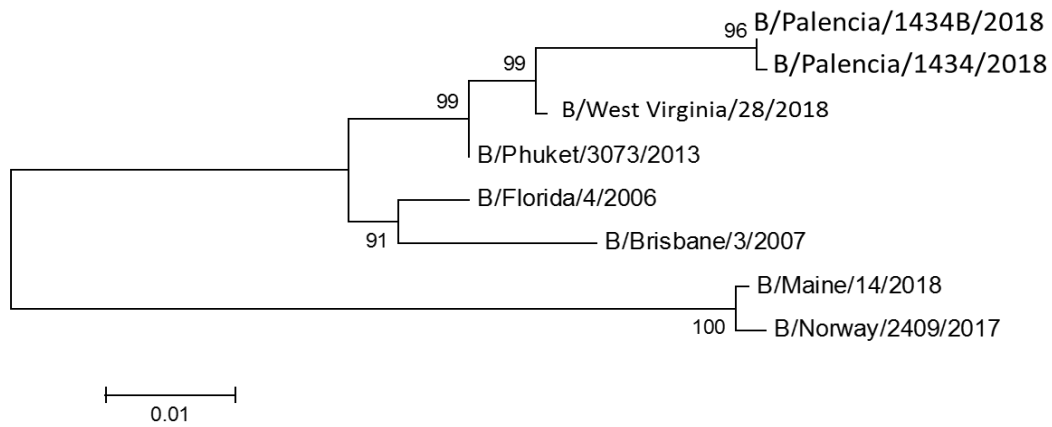
secuencias obtenidas tras la amplificación por PCR multisegmento A no pudo ser obtenido ningún segmento completo.

En el caso del paciente 2 tras el análisis bioinformático de las librerías realizadas con el producto de amplificación de la PCR multisegmento B se obtuvo el genoma completo de una Gripe B linaje Yamagata muy parecido al de la paciente 1, los segmentos NP y NS no pudieron ser analizados filogenéticamente debido a que las secuencias consenso no estaban completas (Figura 41). La secuencia completa de este virus fue depositada en la base de datos GISAID con el nombre de “B/Palencia/1434B/2018” (EPI\_ISL\_358019).

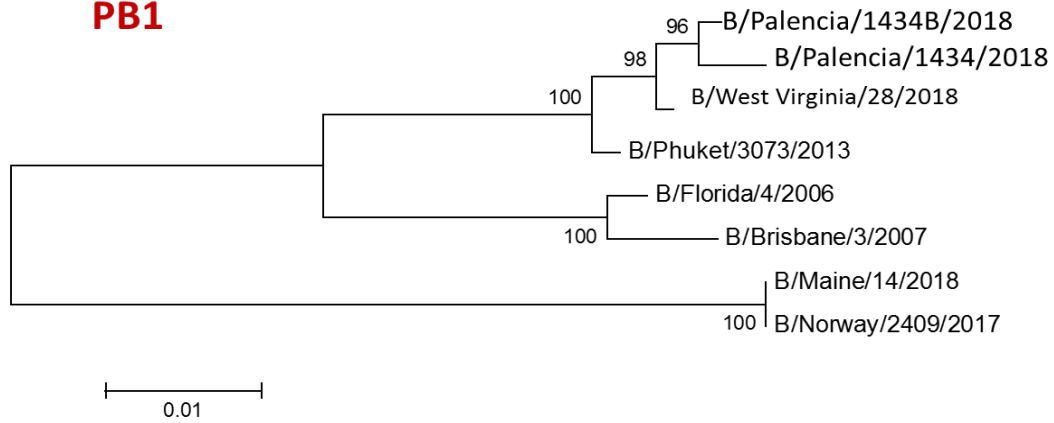
Los resultados obtenidos empleando la PCR multisegmento A mostraban segmentos combinados de Gripe A y B, el análisis clasificó los segmentos PB2, PB1, HA, NP, NA y M pertenientes a Gripe B y los segmentos PA y NS los asoció a Gripe A.

En ninguno de los casos fue posible la obtención de secuencia completa de la gripe A/H3N2 presentes en las muestras de estos pacientes. Para poder caracterizar genéticamente el virus de la gripe A se secuenció un fragmento del gen de la HA mediante Sanger para su posterior análisis por *BLASTn* (Materiales. Apartado 3.5.5). Los virus detectados se asociaron con los virus “B/Phuket/3073/2013” y “A/Colorado/7204/2018”.

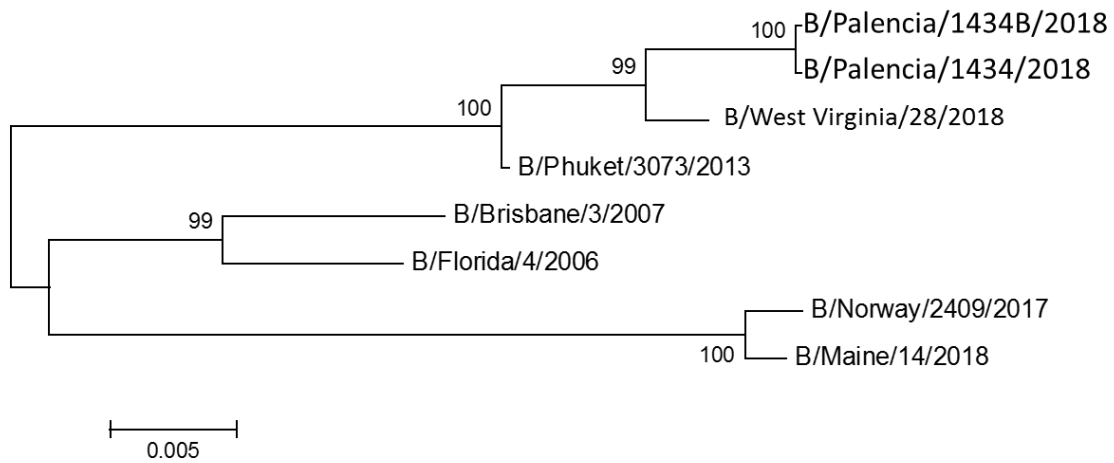
## PB2



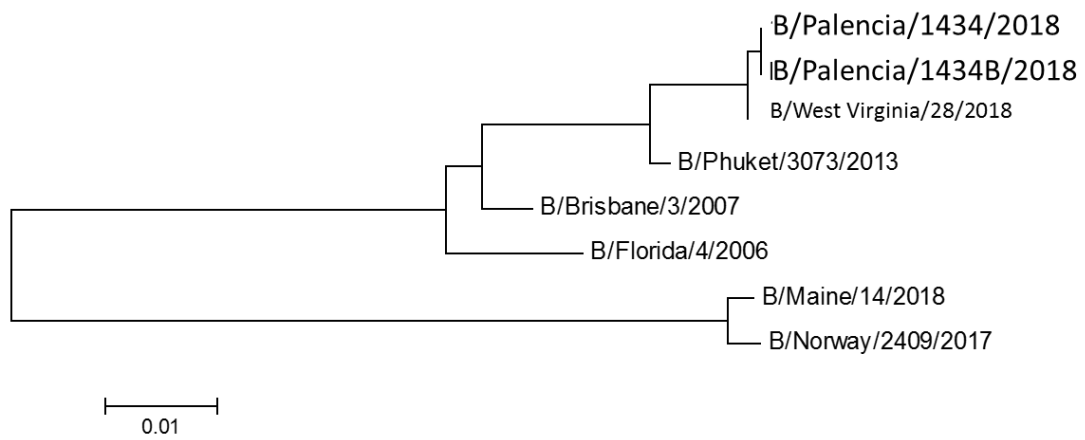
## PB1



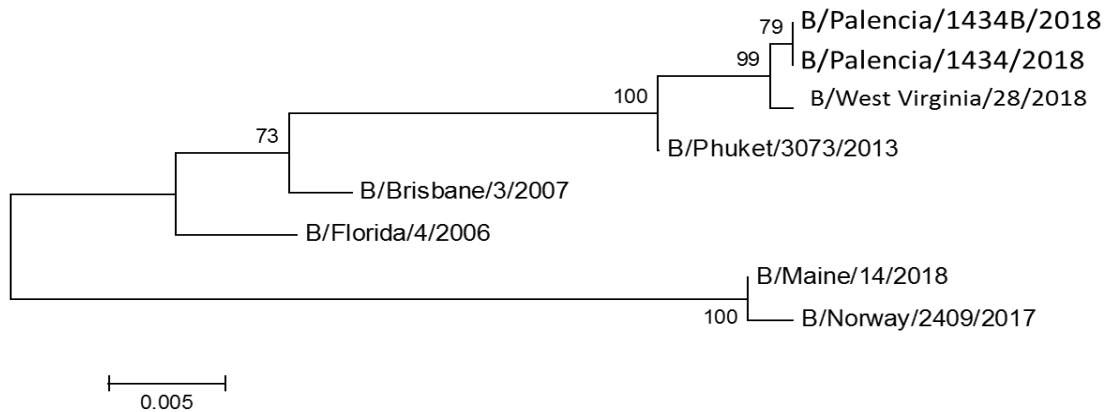
**PA**



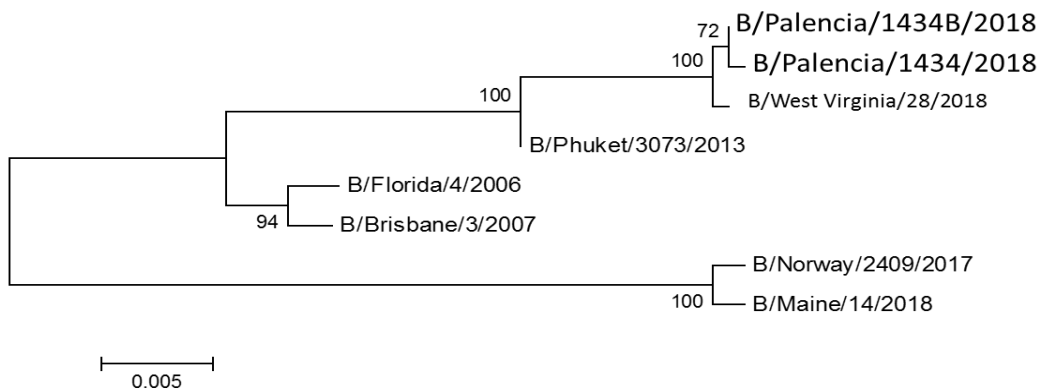
**HA**



**M**



**NA**



**Figura 41:** Análisis filogenético realizado por el modelo Hasegawa-Kishino-Yano del método *Maximum Likelihood* con un respaldo estadístico de 5000 repeticiones. En el que se muestra el análisis filogenético de los segmentos PB2, PB1, PA, HA, M y NA de los virus “B/Palencia/1434/2018” y “B/Palencia/1434B/2018” junto con otros virus de la gripe B.

## DISCUSIÓN

En los últimos años las tecnologías disponibles para WGS del virus de la gripe están siendo cada vez más asequibles y disponibles gracias una constante mejora y evolución de las herramientas empleadas.

El empleo de estas técnicas en el Laboratorio de Referencia de Gripe requiere de la estandarización de unos procedimientos que puedan ser normalizados como son el enriquecimiento viral, las PCRs multisegmento, la elaboración de librerías o el desarrollo de herramientas periféricas como el análisis bioinformático. A pesar de todo, el ajuste de estos

métodos puede ser un reto para un Laboratorio de Referencia de Gripe debido a la diversidad de posibilidades que no siempre están claras en la literatura [20]

La aplicación de técnicas de enriquecimiento es habitual en estudios de metagenómica, especialmente en la búsqueda de nuevos virus, aunque su uso a veces no está justificado ya que este tratamiento no es igual de efectivo en todos los virus. En el desarrollo de esta tesis se estandarizó un método de enriquecimiento por 3 métodos (Métodos. Apartado 4.5.1.) que consistió en centrifugación, filtración y digestión enzimática de la muestra respiratoria similar al considerado por Hall y col. el mejor tratamiento para el enriquecimiento del virus de la gripe [9]. Los resultados obtenidos tras el análisis entre las muestras enriquecidas por el método recomendado y muestras no enriquecidas aportan resultados, tanto en cuantificación por PCR tiempo real como tras la PCR multisegmento, que han demostrado que el protocolo de enriquecimiento de 3 métodos no contribuye a aumentar la proporción de virus de la gripe en la muestra (Figura 29 y Tabla 31. Página 170).

En el caso del virus de la gripe la amplificación exitosa del genoma completo gracias a las PCRs multisegmento parece el método más efectivo para aumentar el número de lecturas de las secuencias del virus de la gripe. La amplificación específica de los 8 segmentos de los virus de la gripe A y B empleando las PCRs multisegmento permite la obtención del genoma completo del virus y sus variantes minoritarias [16-18]. El empleo de la PCR multisegmento universal permite la amplificación del genoma completo de cualquier gripe [18] en una sola PCR, el empleo de este método evitaría la caracterización previa del tipo de virus de la gripe para realizar la PCR multisegmento adecuada y permitiría la detección de coinfecciones [16,17]. A pesar de esta ventaja las PCRs multisegmento A y B son más robustas por el momento que la PCR multisegmento universal (Figura 30. Página 172). El análisis de los virus de la gripe presentes en las muestras respiratorias de los pacientes pediátricos 1 y 2 permitió evaluar las diferencias entre los resultados obtenidos tras realizar la PCR multisegmento A y B. En el paciente 2 empleando la PCR multisegmento A los resultados obtenidos proponían como secuencias consenso segmentos tanto de gripe A como de gripe B. Esto puede ser debido a la existencia de una mayor carga viral de gripe B en los pacientes estudiados, que favorece un sesgo en la amplificación de la PCR multisegmento hacia el tipo de virus más abundante en la muestra. A pesar de esto, la detección de los virus de la gripe de tipo distinto al de la PCR multisegmento empleada en las muestras de



ambos pacientes fue posible probando que el empleo de estas PCRs permite detectar coinfecciones, hecho que hasta este momento no se ha documentado en la bibliografía.

Además, el uso de PCRs de amplificación de genoma completo mejora los resultados de métodos publicados que partían directamente de muestra de pulmón y fueron amplificados aleatoriamente tras el enriquecimiento de la muestra por captura por sondas [151].

Existen diferentes plataformas de secuenciación disponibles con capacidades diferentes como son la plataforma Ion Torrent Personal Genome Machine [242], la plataforma PacBio RS [243], las plataformas Illumina *MiSeq* y *HiSeq* [244] y la nueva plataforma *MinION* [245]. El secuenciador *MinION* es un secuenciador portátil que puede leer secuencias largas [245], su carácter portátil le hace muy interesante para el laboratorio de Referencia de Gripe debido a que por las características del secuenciador puede ser buen candidato para ser empleado en situaciones como pandemias o epidemias en zonas remotas [7]. Por ello fueron comparados los resultados del secuenciador Illumina *MiSeq* disponible en la Unidad de Genómica del CNM, ISCIII, con los obtenidos con esta plataforma. Las secuencias obtenidas de los 9 virus analizados presentaron gran similitud en todos los segmentos del genoma del virus, a pesar de esto en las secuencias obtenidas empleando la plataforma *MinION* fueron detectadas variantes minoritarias de nucleótidos soportadas por decenas o centenares de lecturas que no estaban presentes en las secuencias consenso obtenidas con el secuenciador *MiSeq* (Tabla 32. Página 174), estas variantes minoritarias no fueron detectadas en las secuencias obtenidas por método Sanger y WGS con la plataforma *MiSeq* (Figura 31. Página 175). Resultados similares fueron previamente publicados por trabajos que describen [7,246] su alta tasa de error y una gran variabilidad de resultados entre carreras. Pese a las ventajas que el secuenciador *MinION* puede aportar al laboratorio de Referencia de Gripe la aparición de errores limita su posibilidad de uso en un laboratorio de salud pública ya que estos cambios debido a errores en la secuenciación pueden representar mutaciones tan importantes como las que representan la resistencia a antivirales [247].

La elección de plataforma de análisis bioinformático es muy importante ya que el método de ensamblaje de virus segmentado es determinante para obtener una secuencia completa fiable. En el caso de INSAFLU [25] no pudieron ser obtenidas las secuencias completas en los 3 virus de la gripe B debido a la influencia de la elección de la secuencia de referencia para el ensamblaje introducida por el usuario en esta plataforma (Tabla 33. Página 176) en comparación con IRMA que no requiere la introducción de secuencias de referencia debido a que emplea el método de

ensamblaje por refinamiento iterativo [21]. En el caso del virus “B/CastillaLaMancha/2439/2017” y los diferentes segmentos disponibles según la referencia elegida demuestra que los segmentos de los virus de gripe tienen una historia evolutiva independiente [110], este hecho dificulta aún más la elección de la mejor secuencia de referencia para la obtención de los 8 segmentos del genoma del virus.

Los resultados obtenidos durante la estandarización de estos protocolos prueban la capacidad de este método de detectar virus a baja concentración, identificar múltiples virus de la gripe en un único análisis y obtener un conocimiento funcional del genoma completo de los virus de la gripe.

El análisis bioinformático de las muestras procedentes de la paciente pediátrica 1 detectó una gripe B linaje Yamagata con un segmento de M con muchas variantes minoritarias con respecto la secuencia consenso generada (Figura 37. Página 182). El análisis de las secuencias secundarias mostró un fragmento de 302 nt identificado como M de gripe A.

Con el fin de comprobar si la aparición de tantos cambios y del fragmento de gripe A fue debido a errores generados por el análisis bioinformático se realizó el análisis de las secuencias por ensamblaje *de novo* con la plataforma Virus TAP [24]. Este análisis detectó un fragmento de 602 nt en el segmento M de gripe B filogenéticamente relacionada con el linaje Victoria. Tras dos WGS posteriores realizadas con la misma muestra no se volvió a secuenciar la matriz perteneciente al linaje Victoria.

Teniendo en cuenta que esta matriz del linaje Victoria era similar a las circulantes esa misma temporada podríamos estar ante un caso de contaminación cruzada. La contaminación cruzada es uno de los mayores riesgos de la secuenciación de nueva generación basada en amplicones [23] este protocolo consta de diferentes puntos que puede ser susceptible de contaminación cruzada como pueden ser los reactivos, el equipo [29] o en la preparación de la PCR de indexado de las librerías [248] por ello las mutaciones, deleciones, inserciones y otros eventos detectados por secuenciación de nueva generación deben ser analizados con detalle debido a la sensibilidad de estas técnicas y el riesgo de amplificar errores [23].

Otra de las posibles razones para la detección del fragmento puede ser diferencia de cobertura del segmento secuenciado, el análisis por ensamblaje *de novo* puede ser problemático debido a

que la síntesis *de novo* solo produce genomas únicos de gripe el 21% de las veces debido al carácter segmentado de su genoma [21].

El análisis de las secuencias demostró la presencia de coinfección en las muestras de los pacientes 1 y 2 y la similitud en los segmentos de los virus de la gripe B linaje Yamagata de ambos hermanos, no siendo posible analizar la similitud en gripe A debido a que ningún virus pudo ser secuenciado de manera completa.

La similitud del virus de gripe B en ambos hermanos en los ocho segmentos y la ausencia de cambios conocidos que afectasen a la virulencia o la resistencia a antivirales obtenidos gracias a estas técnicas aportan información valiosa en este caso.

En resumen, la aplicación de estas técnicas aporta información rápida y completa sobre el virus y su diversidad siendo necesario aún una caracterización previa para aplicar los métodos más adecuados y la comprobación de eventos importantes observados en la muestra. La puesta a punto de diferentes técnicas para resolver las dificultades de esta técnica unida al conocimiento del personal de laboratorio y las técnicas complementarias disponibles la hacen una técnica que aporta gran información en los laboratorios de salud pública.

## **5.4. OBJETIVO IV: Estudio de virus respiratorios ADN de gran tamaño y complejidad detectados en hospedadores no humanos: nuevos adenovirus en murciélagos ibéricos**

Estudios previos realizados por nuestro grupo de investigación caracterizaron la presencia de adenovirus en la Península Ibérica [249]. La disponibilidad de las muestras biológicas tomadas en murciélagos y previamente estudiadas que resultaron positivas a AdV, permitió la realización de ajustes de métodos para la secuenciación de genoma completo de virus ADN de gran tamaño.

En este objetivo fue analizada la presencia de adenovirus en 1392 exudados orofaríngeos y 325 muestras de heces de 27 especies de murciélagos diferentes (Tabla 10. Página 75) obtenidos durante el desarrollo de los proyectos de investigación VIROBAT (SAF2006-12784-C02, SAF2009-09172, SAF2013-47149P y SAF2017-89355-P). Los adenovirus fueron detectados en 12 especies de las familias Vespertilionidae y Rhinolophidae.

A estas muestras se les realizaron dos PCR convencionales para la caracterización la presencia de AdVs en la Península Ibérica. La PCR ADN polimerasa para adenovirus (Métodos. Apartado 4.13.4) consistía en la amplificación de un fragmento del gen de la ADN-dependiente de ADN polimerasa, ampliamente utilizado para taxonomía de Mastadenovirus y muy conservado. La PCR de hexón para adenovirus (Métodos. Apartado 4.13.3) amplifica un fragmento del gen que codifica para la proteína del hexón que es un gen mucho más variable. El estudio de ambos genes tiene como finalidad la obtención de dos filogenias que pueden ser comparadas aportando información sobre las características genéticas de los adenovirus de murciélagos y su relación evolutiva [249].

El producto de amplificación fue visualizado por electroforesis en gel de agarosa al 2% teñidos con *GelRed* (Materiales. Apartado 3.5.1). Los alineamientos fueron realizados con el programa *CLUSTAL X 2.0* y los análisis filogenéticos con el programa *MEGA 5.2*.

Los resultados obtenidos fueron 50 muestras respiratorias y 26 heces positivas. Los análisis de fragmento del gen del hexón y de la polimerasa ADN dependiente de ADN indicaron que todos estos virus de murciélagos pertenecían al género Mastadenovirus. La caracterización de estos AdVs de murciélago mostró que la diversidad de mastadenovirus de murciélagos es dependiente de hospedador.

La disponibilidad de muestras respiratorias y de heces de los murciélagos con AdVs caracterizados permitió la estandarización de las técnicas WGS de AdV desde muestra directa.

Las muestras empleadas fueron 8 EOR pertenecientes a 3 *Pipistrellus kuhlii* y 5 *Pipistrellus pygmaeus* (Tablas 12) que previamente resultaron positivas para AdV. Estas muestras fueron sometidas a dos flujos de trabajo con métodos de enriquecimiento viral diferentes para realizar el estudio comparativo del mejor método para WGS de AdV. Por un lado estas muestras fueron enriquecidas viralmente por 3 métodos (Métodos. Apartado 4.5.1) y los extractos posteriormente amplificados por SISPA ADN (Métodos. Apartado 4.5.2.1), las librerías fueron realizadas con el kit *Nextera XT* (Métodos. Apartado 4.8.) y secuenciadas por la plataforma *MiSeq* (Materiales. Apartado 3.5.4). El otro método empleado fue el empleo del método SISPA ADN en los extractos procedentes de los EOR y la preparación de librerías con el kit *Kapa Hyper Plus* enriquecidas con el método de captura por sondas empleando el kit *HyperCap Target Enrichment* (Métodos. Apartado 4.9) para la secuenciación en la plataforma *MiSeq*.

El análisis bioinformático de las secuencias fue realizado como se describe en el Apartado 4.10

de Métodos empleando la plataforma *Virus TAP* [24] y las herramientas de análisis empleadas en USAMRIID.

Los resultados obtenidos mostraron que el enriquecimiento con sondas permitió la secuenciación de 16884 nt del genoma de AdV directamente de muestra de respiratoria.

#### 5.4.1 Estudio de la presencia de AdV en las muestras de murciélagos ibéricos

En el desarrollo del estudio de caracterización de AdV de murciélagos fueron analizadas un total de 1717 muestras, de las cuales 1392 fueron EOR y 325 heces. Estas muestras fueron obtenidas de 27 especies de murciélagos de las 32 existentes en Europa (<http://secemu.org>). De esas 27 especies analizadas 22 pertenecían a la familia Vespertilionidae, 1 a la familia Miniopteridae y 4 especies a la familia Rhinolophidae.

De los análisis realizados por las PCRs convencionales de hexón y ADN polimerasa (Métodos. Apartado 4.13.3 y 4.13.4) se obtuvieron 50 muestras positivas (3,6 %) en los EOR y 26 muestras positivas (8,3%) en las heces.

Del total de individuos analizados en 70 de ellos se detectó ADN de AdV y en 3 de estos murciélagos fue posible la detección del virus tanto en EOR como en heces. La amplificación exitosa del fragmento parcial del hexón se obtuvo en 69 muestras (49 EOR y 20 heces) y para el fragmento de la polimerasa en 35 muestras (14 EOR y 21 heces). Todos los fragmentos amplificados fueron confirmados por secuenciación y depositados en la base de datos *GenBank* (Materiales. Apartado 3.5.6) (Tabla 37).

Andalucía fue la región con mayor distribución de mastadenovirus de murciélagos incluyendo varios géneros de la familia Vespertilionidae (*Pipistrellus*, *Myotis* y *Nyctalus*) y Rhinolophidae (*Rhinolophus*). La mayoría de los murciélagos del género *Pipistrellus* fueron muestreados en Andalucía debido a esto la mayoría de AdV detectados se obtuvieron en murciélagos del sur de España. Sin embargo, en el caso de los murciélagos del género *Rhinolophus* fueron muestreados 59 murciélagos procedentes del sur de España de los cuales 10 muestras resultaron positivas para AdV mientras que de los 78 murciélagos del género *Rhinolophus* muestreados en el País Vasco no se detectó ningún positivo.

Especie de murciélago			EOR	H	Secuencias hexón	Secuencias ADN polimerasa
Familia	Nombre	Abv				
Vespertilionidae	<i>Barbastella barbastellus</i>	Bba	0/38	0/4	N/A	N/A
	<i>Eptesicus isabellinus</i>	Eis	0	0/8	N/A	N/A
	<i>Eptesicus serotinus</i>	Ese	0	0/14	N/A	N/A
	<i>Hypsugo savii</i>	Hsa	0/31	3/26	HM856338,41,42	JX065121, 22, MG208122
	<i>Myotis alcaethoe</i>	Mal	0	0/1	N/A	N/A
	<i>Myotis bechsteinii</i>	Mbe	1/18	0/2	MF540611	N/A
	<i>Myotis blythii</i>	Mbl	0/29	0	N/A	N/A
	<i>Myotis capaccinii</i>	Mca	0/15	0	N/A	N/A
	<i>Myotis daubentonii</i>	Mda	0/63	0/41	N/A	N/A
	<i>Myotis emarginatus</i>	Mem	3/56	0	MF540608-10	N/A
	<i>Myotis escalerae</i>	Mes	0/13	0	N/A	N/A
	<i>Myotis myotis</i>	Mmy	1/79	0/1	HM856353	N/A
	<i>Myotis mystacinus</i>	Mmt	0/2	0/8	N/A	N/A
	<i>Myotis nattereri</i>	Mna	0/36	0/3	N/A	N/A
	<i>Nyctalus noctula</i>	Nno	3/122	0	MF540597-99	N/A
	<i>Nyctalus lasiopterus</i>	Nlas	10/139	6/40	HM856327-34,39-40,43,45-47, 50, MG132211	JX065117-20, 23,25-26,28
	<i>Nyctalus leisleri</i>	Nle	1/19	3/26	HM856344,48, 51-52	JX065124,27,29
	<i>Pipistrellus kuhlii</i>	Pku	12/350	2/4	MF540577-85,87,89	MF404970-73,75,86
	<i>Pipistrellus pipistrellus</i>	Ppi	0/29	0/4	HM856349	N/A
	<i>Pipistrellus pygmaeus</i>	Ppy	6/36	11/120	MF540575-76, 86,88,90-96	MF404968-69,74, 76-79, 80-85,87-89
	<i>Plecotus auritus</i>	Pau	0/11	0/8	N/A	N/A
	<i>Plecotus austriacus</i>	Pas	0/10	0/6	N/A	N/A
Miniopteridae	<i>Miniopterus schreibersii</i>	Msc	0/152	0/2	N/A	N/A
Rhinolophidae	<i>Rhinolophus euryale</i>	Reu	6/49	0	MF540600-02,12-13 HM856335	N/A
	<i>Rhinolophus ferrumequinum</i>	Rfe	7/90	1/3	MF540603-07,14 HM856336-37	N/A
	<i>Rhinolophus hipposideros</i>	Rhi	0/4	0/4	N/A	N/A
	<i>Rhinolophus mehelyi</i>	Rme	0/1	0	N/A	N/A
Total	27/32		50/1392	26/325	69 49EOR + 20H	35 14EOR + 21H

**Tabla 37:** Especies de murciélago estudiadas, resultados del estudio de AdV en EOR y H, año de captura y número de identificación de *GenBank*

Las tres especies de *Nyctalus* (*N. noctula*, *N. lasiopterus* y *N. leisleri*, que aportaron 23 AdVs de murciélago detectados) y dos de las tres de *Pipistrellus* (*P. kuhlii* y *P. pygmaeus*, que aportaron 28 AdVs de murciélago detectados) contribuyeron con el mayor número de muestras positivas. De acuerdo al género *Rhinolophus* en 2 de las 4 especies (*R. euryales* y *R. ferrumequinum*, donde se detectaron 14 AdV de murciélagos) solo pudieron detectarse los AdV en EOR.

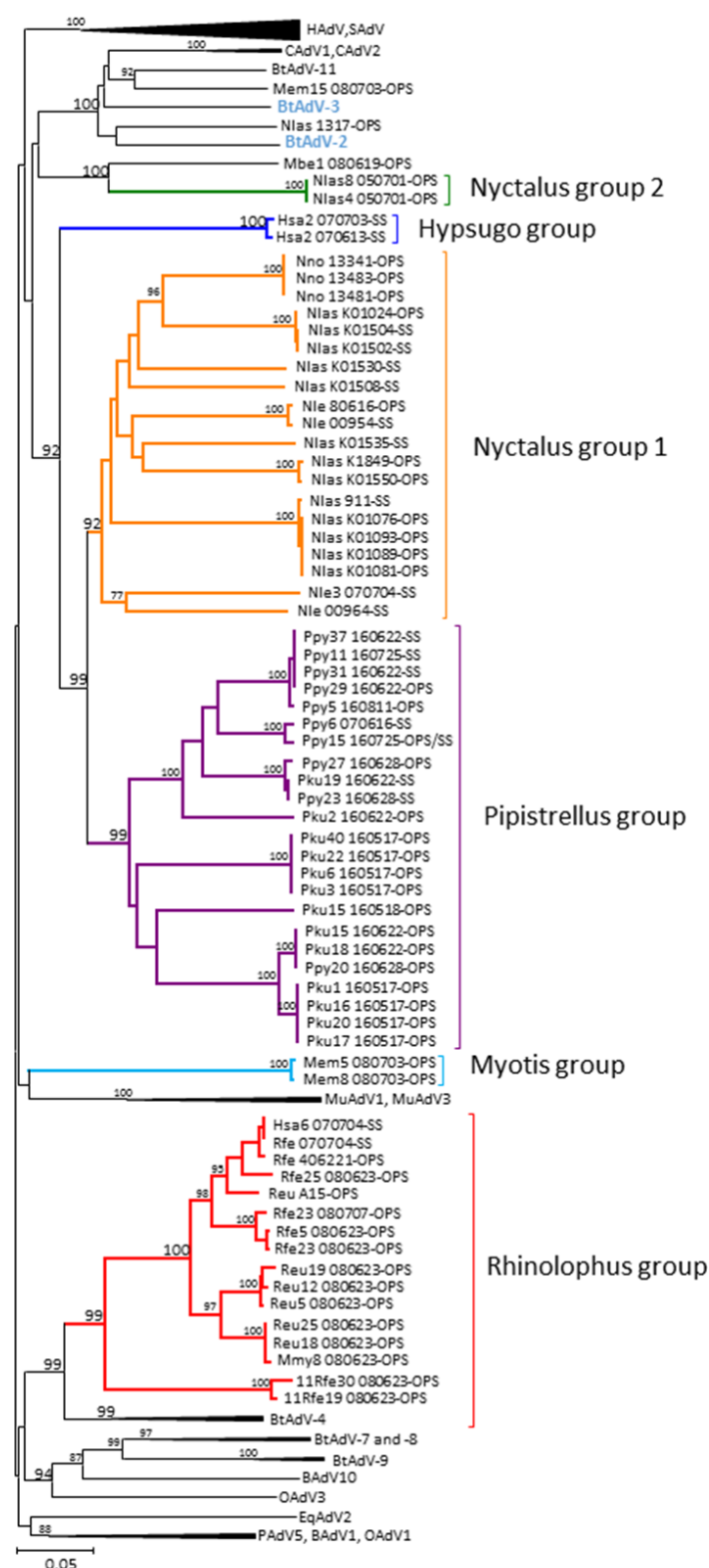
#### 5.4.2 Análisis de las secuencias de hexón de AdV de murciélagos

El análisis filogenético de los fragmentos de ~768 nt obtenidos tras la PCR convencional de hexón para AdV (Métodos. Apartado 4.13.3) mostró que estos virus detectados formaban parte del género *Mastadenovirus* y que además los virus detectados en murciélagos de la familia Rhinolophidae y Vespertilionidae se diferenciaban en dos grupos filogenéticos bien respaldados estadísticamente.

Este análisis del fragmento del gen del hexón se realizó con el programa *MEGA 5.2* empleando el método *Neighbor-Joining* con un método *Tamura 3 parámetros* con un respaldo estadístico de 5000 repeticiones (Métodos. Apartado 4.15.3) (Figura 42).

En este análisis se observó la relación de 2 AdV de murciélagos detectados en EOR de las especies *Nyctalus lasiopterus* y *Myotis emarginatus* con el mastadenovirus G de murciélago (KX871230) procedente de la especie *Corynorhinus rafinesquii* capturado en EEUU [146]. A su vez, estos 3 murciélagos se agruparon con las cepas de referencia AdV 2 (mastadenovirus A de murciélago) (JN252129) y 3 (mastadenovirus B de murciélago) (GU226970) de murciélago, detectados en murciélagos de la especie *Pipistrellus pipistrellus* [137] y *Myotis ricketti* [139], respectivamente.

El resto de las 73 secuencias de AdV de murciélagos analizadas se agruparon con un respaldo estadístico alto en 6 grupos basados en la especie de murciélago en el que fueron detectados. Estos nuevos grupos diferenciados por hospedador se denominaron grupo de *Pipistrellus*, el grupo de *Nyctalus* 1, el grupo *Nyctalus* 2, grupo de *Hypsugo* y el grupo de *Myotis* dentro de la familia Vespertilionidae y el grupo de *Rhinolophus* de la familia Rhinolophidae (Figura 42).



**Figura 42:** Árbol filogenético basado en el análisis del fragmento del gen de hexón. El árbol fue estimado con el programa *MEGA 5.2* usando el método *Neighbor-Joining* con el modelo *Tamura 3 parámetros*. El soporte estadístico fue de 5000 repeticiones.



El grupo 1 de *Nyctalus* se conformó con un grupo de 13 AdVs detectados en la especie *N. lasiopterus*, 4 detectados en *N. leisleri* y 3 detectados en *N. Noctula*, este grupo se relacionó con el grupo *Pipistrellus*. En el grupo 2 de *Nyctalus*, 2 AdVs detectados en la especie *N. lasiopterus* formaron un grupo diferente al del resto de *Nyctalus* junto con un AdV detectado en un murciélago de la especie *Myotis bechsteinii*.

El grupo *Pipistrellus* fue un grupo bien definido estadísticamente formado por 13 AdVs detectados en murciélagos de la especie *P. kuhlii* y 9 AdVs detectados en la especie *P. pygmaeus*. Dos grupos pequeños pero bien respaldados estadísticamente formaban parte de este gran grupo de AdVs detectados en murciélagos de la familia Vespertilionidae: el grupo *Hypsugo* y el grupo *Myotis*.

El grupo *Hypsugo* estaba relacionado con el grupo *Nyctalus* 1 y el grupo *Pipistrellus* y estaba conformado por 2 AdVs detectados en murciélagos de la especie *Hypsugo savii*.

Dos AdVs detectados en murciélagos de la especie *Myotis emarginatus* constituyeron el grupo *Myotis*.

El grupo *Rhinolophus* formaba un grupo bien respaldado estadísticamente diferenciado de los virus detectados en muestras de murciélagos de la familia Vespertilionidae. Este grupo estaba formado por 8 AdVs detectados en murciélagos de la especie *R. ferrumequinum* y 6 AdVs de la especie *R. euryales*. Una de las particularidades de este grupo es la asociación de un AdV obtenido en un murciélago *Myotis myotis* y de otro AdV detectado en un murciélago de la especie *Hypsugo savii* con este grupo. Tres AdVs (KT698853, NC029899, NC029902) procedentes de otro análisis en China realizado en murciélagos de la especie *R. sinicus* [133] también formaron parte de este grupo *Rhinolophus*.

#### 5.4.3 Análisis de las secuencias de ADN polimerasa de AdV de murciélagos

El análisis filogenético de los fragmentos de ~450 nt obtenidos tras la PCR convencional de ADN polimerasa para AdV (Métodos. Apartado 4.13.4) mostraron que los grupos definidos en este gen estaban claramente asociados por el huésped aunque con menor soporte y resolución en algunos nodos en comparación con el análisis filogenético del fragmento del gen de hexón.

Este análisis del fragmento del gen de la ADN polimerasa se realizó con el programa *MEGA 5.2* empleando el método *Neighbor-Joining* con un método *Kimura 2 parámetros* con un respaldo estadístico de 5000 repeticiones (Métodos. Apartado 4.15.4) (Figura 43).

Tras el análisis se agruparon 5 AdV detectados en murciélagos *Pipistrellus pygmaeus* junto con la cepa de referencia AdV 2 (mastadenovirus A de murciélago) (JN252129) de murciélago, detectados en murciélagos de la especie *Pipistrellus pipistrellus* [137], en un grupo que incluía 3 AdVs detectados en murciélagos *Pipistrellus kuhlii* (MF404975), *N. lasiopterus* (JX065123) y *P. pygmaeus* (KM43091) capturados en Hungría [14]. Este grupo, que incluía 6 AdV detectados en murciélagos del género *Pipistrellus*, se agrupó por separado del resto de AdV detectados en murciélagos del género *Pipistrellus*.

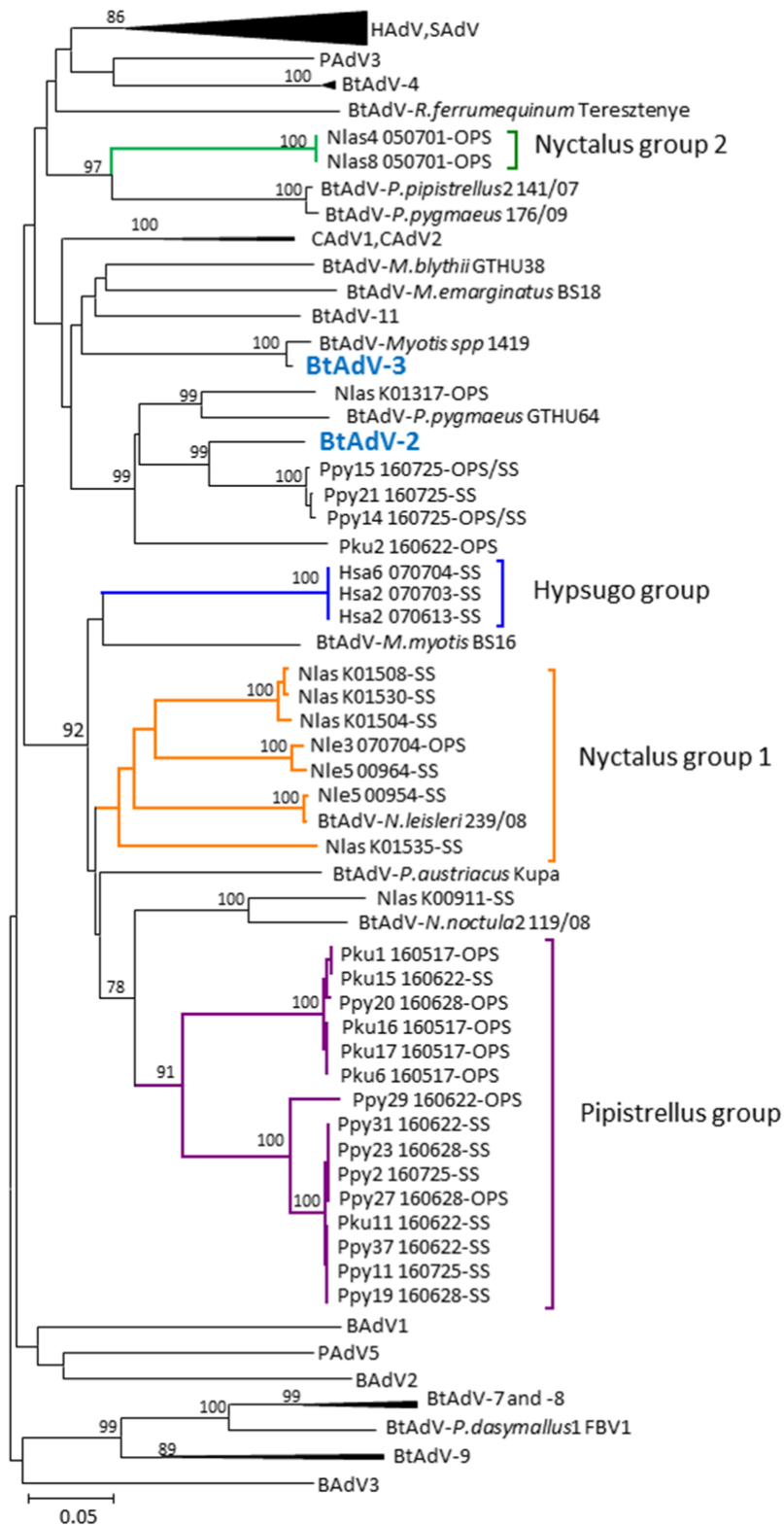
El análisis de las secuencias obtenidas del género *Nyctalus* mostró la formación de un grupo similar al grupo *Nyctalus* 1 descrito en el análisis filogenético del fragmento del gen de hexón, con la excepción de que 1 AdV detectado en un murciélago *N. lasiopterus* del grupo 1 conformó un nuevo grupo junto a un AdV detectado en un *N. noctula* (KM043110) capturado en Hungría. El grupo *Nyctalus* 2 agrupó con dos AdVs detectados en murciélagos *P. Pipistrellus* (KM043096) y *P. pygmaeus* (KM43091) capturados en Hungría.

El grupo *Pipistrellus* estaba formado por 5 AdV detectados en murciélagos de la especie *P. kuhlii* y 12 murciélagos de la especie *P. pygmaeus* resultando un grupo similar al observado con el análisis parcial del gen hexón.

En el grupo *Hypsugo* un AdV detectado en 1 murciélago *H. savii* (MG208122, que previamente en el análisis del fragmento de gen de hexón agrupó en el grupo *Rhinolophus*, tras el análisis del fragmento del gen de la ADN polimerasa agrupó en el grupo *Hypsugo* con el resto de AdVs procedentes de los murciélagos de este género.

En este análisis no se obtuvieron resultados positivos para los murciélagos del género *Rhinolophus*.

Los valores del análisis de distancias por pares realizados con las secuencia de aminoácidos parciales obtenidas de la PCR de ADN polimerasa de AdV respaldaron la formación de estos nuevos grupos por hospedador (Tabla 38).



**Figura 43:** Árbol filogenético basado en el análisis del fragmento del gen de la ADN polimerasa. El árbol fue estimado con el programa *MEGA 5.2* usando el método *Neighbor-Joining* con el modelo *Kimura 2 parámetros*. El soporte estadístico fue de 5000 repeticiones.

Grupo de AdV	Nombre provisional de los virus	Abv.	% distancia entre pares	
			BtAdV-3	BtAdV-2
AdVs asociados con BtAdV-2	Mastadenovirus <i>P. pygmaeus</i> 14 160725	Ppy14 160725	31.4	11.8
	Mastadenovirus <i>P. pygmaeus</i> 15 160725	Ppy15 160725	31.4	12.7
	Mastadenovirus <i>P. pygmaeus</i> 21 160725	Ppy21 160725	32	12.2
Potenciales nuevos AdVs de murciélago	Mastadenovirus <i>N. lasiopterus</i> K01317	Nlas K01317	33.8	25.5
	Mastadenovirus <i>N. lasiopterus</i> K01508	Nlas K01508	35	41.8
	Mastadenovirus <i>N. lasiopterus</i> K01530	Nlas K01530	35.7	42.5
	Mastadenovirus <i>N. lasiopterus</i> K01504	Nlas K01504	35.7	42.5
	Mastadenovirus <i>N. leisleri</i> 3 070704	Nle3 070704	32.5	42.5
	Mastadenovirus <i>N. leisleri</i> 00964	Nle 00964	34.5	44
	Mastadenovirus <i>N. leisleri</i> 5 00954	Nle5 00954	39.7	41.8
	Mastadenovirus <i>N. lasiopterus</i> K01535	Nlas K01535	41.8	52.1
	Mastadenovirus <i>N. lasiopterus</i> K00911	Nlas K00911	43.5	46.1
	Mastadenovirus <i>N. lasiopterus</i> 4 050701	Nlas4 050701	42.6	38.8
	Mastadenovirus <i>N. lasiopterus</i> 8 050701	Nlas8 050701	42.6	38.8
	Mastadenovirus <i>P. kuhlii</i> 2 160622	Pku2 160622	37.5	23.8
	Mastadenovirus <i>P. kuhlii</i> 1 160517	Pku1 160517	39.4	39
	Mastadenovirus <i>P. kuhlii</i> 15 160622	Pku15 160622	39.4	39
	Mastadenovirus <i>P. pygmaeus</i> 20 160628	Ppy20 160628	38.7	39
	Mastadenovirus <i>P. kuhlii</i> 16 160517	Pku16 160517	38.7	38.3
	Mastadenovirus <i>P. kuhlii</i> 17 160517	Pku17 160517	38.7	38.3
	Mastadenovirus <i>P. kuhlii</i> 6 160517	Pku6 160517	38.7	38.3
	Mastadenovirus <i>P. pygmaeus</i> 29 160622	Ppy29 160622	44.8	36.5
	Mastadenovirus <i>P. pygmaeus</i> 31 160622	Ppy31 160622	41.2	41.6
	Mastadenovirus <i>P. pygmaeus</i> 23 160628	Ppy23 160628	41.2	41.6
	Mastadenovirus <i>P. pygmaeus</i> 2 160725	Ppy2 160725	41.2	41.6
	Mastadenovirus <i>P. pygmaeus</i> 27 160628	Ppy27 160628	41.2	41.6
	Mastadenovirus <i>P. kuhlii</i> 11 160622	Pku11 160622	42	40.9
	Mastadenovirus <i>P. pygmaeus</i> 37 160622	Ppy37 160622	42	40.9
	Mastadenovirus <i>P. pygmaeus</i> 11 160725	Ppy11 160725	42	40.9
	Mastadenovirus <i>P. pygmaeus</i> 19 160628	Ppy19 160628	42	40.9
	Mastadenovirus <i>H. savii</i> 6 070704	Hsa6 070704	41	41.7
	Mastadenovirus <i>H. savii</i> 2 070613	Hsa2 070613	41	41.7
	Mastadenovirus <i>H. savii</i> 2 070703	Hsa2 070703	41	41.7

**Tabla 38:** Clasificación de los mastadenovirus de murciélago tras el análisis de distancia entre pares de la secuencia de aminoácidos obtenida de las secuencias del fragmento de ADN polimerasa.

Según los datos del análisis de distancia por pares, ninguno de los AdV detectados procedentes de murciélagos ibéricos se relacionaron con la especie Bat mastadenovirus A. Del mismo modo, 3 AdV detectados en *Pipistrellus pygmaeus* (Ppy14 160725, Ppy15 160725 y Ppy21 160725) mostraron una distancia por pares de aminoácidos del 12% con el Bat mastadenovirus B, siendo

el umbral de demarcación de nueva especie valores mayores del 15% en el análisis de distancia por pares podríamos decir que son virus similares.

#### 5.4.4 Estudio de los métodos de enriquecimiento para WGS de adenovirus

Tras la detección de AdV en las muestras de murciélagos empleamos 8 EOR pertenecientes a 3 *Pipistrellus kuhlii* y 5 *Pipistrellus pygmaeus* (Tabla 11. Página 75) (Materiales.Apartado 3.3.6) que previamente resultaron positivas para AdV para el estudio comparativo de métodos de enriquecimiento con el fin de obtener por NGS el genoma completo de AdV.

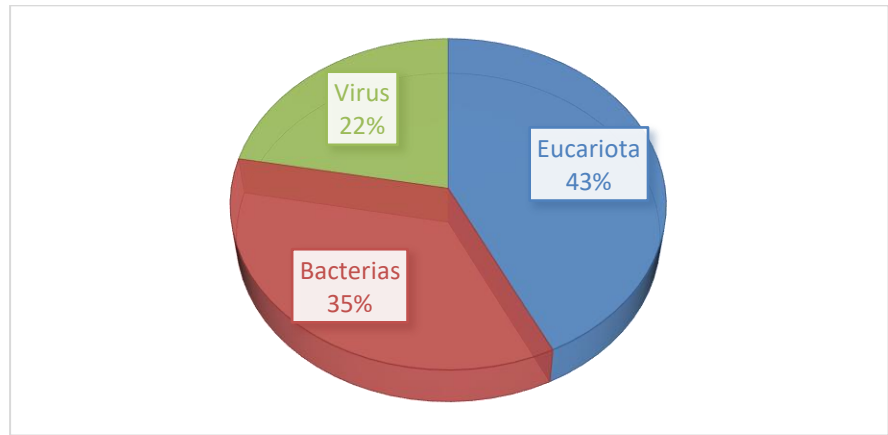
Inicialmente 7 EOR fueron sometidos al flujo de trabajo para el estudio de viroma (Métodos.Apartado 4.5) descrito en este trabajo en el que se les realizó el enriquecimiento viral por 3 métodos, la amplificación por el método SISPA ADN y la preparación de librerías con el kit *Nextera XT*. Por otro lado las mismas muestras fueron amplificadas por el método SISPA ADN y se prepararon las librerías empleando el kit *HyperCap Hyper Plus*, estas librerías fueron enriquecidas mediante captura por sondas empleando el kit *HyperCap Target Enrichment* (Métodos. Apartado 4.9).

Los resultados obtenidos tras el análisis bioinformático tras el enriquecimiento por 3 métodos y la preparación de librerías por *Nextera XT* mostraron que las lecturas obtenidas fueron 14, 5 millones de lecturas de las cuales 6,2 millones son fragmentos de secuencias de eucariotas (43%), 5, 1 millones son fragmentos de secuencias de bacterias (35%) y 3,2 millones fueron fragmentos de secuencias virales (22%) (Figura 44), dentro de las lecturas virales obtenidas se obtuvieron de 4100 fragmentos de AdV que correspondían al 12,5% del total de las secuencias virales.

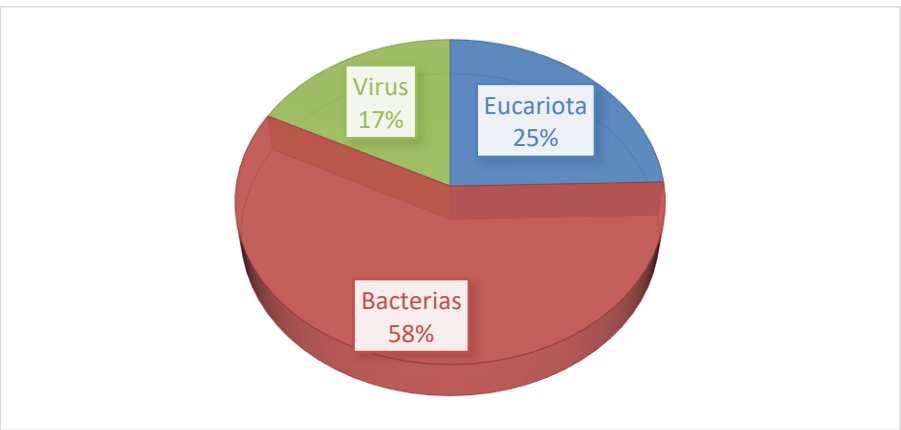
En el caso del análisis bioinformático tras el método de enriquecimiento mediante captura por sondas obtuvimos 18, 2 millones de lecturas procedentes de las 7 muestras. De estos 18, 2 millones de lecturas 4, 5 millones fueron taxonómicamente identificadas como fragmentos de secuencias eucariotas (25%), 10, 6 millones de lecturas fueron fragmentos bacterianos (58%) y 3,2 millones fueron fragmentos virales (17%) (Figura 45). Dentro de los 3,2 millones de lecturas virales se obtuvieron 9000 fragmentos de AdV, representado el 30% de los fragmentos virales obtenidos.

El tamaño de las secuencias obtenidas también fue diferente entre los virus secuenciados empleando el enriquecimiento por 3 métodos con la preparación de librerías por *Nextera XT* y

los virus secuenciados empleando el kit de preparación de librerías *Kapa Hyper Plus* con el enriquecimiento por sondas con el kit *HyperCap Target Enrichment*. El tamaño de secuencia más largo fue obtenido fue el de la secuencia de 2188 nt del AdV “Pku15 160518” (Tabla 39).



**Figura 44:** Distribución del número de fragmentos de genoma obtenidos en los 7 EOR empleados tras el enriquecimiento por 3 métodos y la preparación de librerías con el kit *Nextera XT*.



**Figura 45:** Distribución del número de fragmentos de secuencias obtenidos en los 7 EOR empleados tras el enriquecimiento mediante captura por sondas y la preparación de librerías con el kit *HyperCap Target Enrichment*.

	Enriquecimiento por 3 métodos	Captura por sondas
Pku2 160622	121	359
Pku15 160518	905	2188
Pku40 160517	N/A	113
Ppy14 160725	N/A	N/A
Ppy15 160622	N/A	N/A
Ppy23 160628	176	612
Ppy29 160622	N/A	N/A

**Tabla 39:** Tamaño de secuencia obtenido tras el análisis bioinformático de las secuencias obtenidas tras la aplicación de los dos métodos de enriquecimiento de secuencias comparados

En un análisis posterior se secuenciaron empleando el enriquecimiento mediante captura por sondas dos muestras respiratorias de murciélago positivas para AdV, en las muestras empleadas fueron detectados los AdV “Ppy15 181106” y “Pku 15 150518” (Tabla 11. Página 75). Estas muestras fueron elegidas para realizar la WGS de AdV debido a que tras el análisis con la PCR múltiple tiempo real BRQ presentaron el umbral de ciclo más bajo, “Ppy15 181106” presentó un umbral de ciclo de 23,2 y “Pku15 150518” de 25,7.

El análisis bioinformático de las secuencias obtenidas mostraron que el tamaño obtenido de la secuencia del AdV “Pku15 150518” fue de 4380 nt y de 16884 nt en el caso del AdV “Ppy15 181106” siendo el mayor tamaño de secuencia obtenido de muestra respiratoria directa hasta el momento (Figura 46). La secuencia obtenida está depositada en *GenBank* pendiente de asignación de identificador.



**Figura 46:** Alineamiento realizado con el programa MACAW entre las secuencias del mastadenovirus de murciélago de referencia BtAdV-2 y los AdV Pku15 150518 y Ppy15 181106 obtenidas tras el enriquecimiento mediante captura por sondas.

#### 5.4.5 Infección de adenovirus de murciélago en células VERO E-6

Debido a que no fue posible obtener secuencia completa directamente de la muestra y dado que trabajos previos [133,142,146] habían conseguido infectar con AdV de murciélago células Vero E-6 se realizó una prueba para intentar aumentar la cantidad de virus disponible para su posterior secuenciación.

Las células VERO E-6 (Materiales. Apartado 3.1) fueron infectadas con los AdV de murciélago “Pku15 160518” y “Pku2 160622” siguiendo el protocolo descrito en el apartado 4.2 de Métodos.

Las pruebas realizadas fueron la inoculación de las monocapas de VERO E-6 con 100 y 200 µl de la muestra de campo a 32°C, temperatura corporal de los murciélagos, y a 37°C que es la temperatura a la que crecen las monocapas celulares.

A pesar de la prueba con diferentes concentraciones de inóculo y las diferentes temperaturas de adsorción empleadas con el fin del permitir el crecimiento del virus no fue posible observar ni efecto citopático ni recuperar virus.

## DISCUSIÓN

En esta tesis han sido descritas la detección y las relaciones filogenéticas existentes entre mastadenovirus de murciélagos potencialmente nuevos y los AdV de murciélagos previamente descritos usando dos genes parciales diferentes. Nuestro estudio muestra, por primera vez, la diversidad de los AdVs detectados en murciélagos capturados en el sur de Europa, una región de crucial importancia debido a su ubicación geográfica estratégica como corredor entre África y Europa.

Trabajos previos han demostrado la gran diversidad de AdV encontrados en especies de murciélagos en Europa, Asia y África [136,137,140,142,144,146]. En este estudio de AdV en murciélagos, se examinaron 27 de las 32 especies de murciélagos ibéricos existentes obteniendo resultados positivos en 12 especies de 6 géneros de murciélagos. Un trabajo previo en el centro de Europa, Hungría y Alemania, obtuvo resultados positivos para AdV en 9 especies de las 12 [141] especies con resultados positivos descritas en esta tesis. Con el objetivo de tener una representación amplia de la diversidad de AdV en los murciélagos ibéricos se analizaron un total de 1717 muestras biológicas siendo este trabajo el estudio de caracterización de AdV de murciélago con mayor número de muestras realizado. Estos murciélagos fueron capturados dentro de España en una variedad de hábitats que incluye desde las cadenas montañosas de los Pirineos y el Cantábrico en el norte hasta el Mediterráneo , frontera natural con África en la que se ha descrito la existencia de un flujo genético en varias especies de murciélagos a través del Estrecho de Gibraltar [250].

El porcentaje de AdV positivos en los murciélagos analizados fue del 3.6% en EOR y 8.3% en H. Un estudio previo realizado en el centro de Europa obtuvo un porcentaje de AdVs positivos del 18.6% en muestras obtenidas en Alemania y del 9.9% en las muestras obtenidas en Hungría [141], estas diferencia de porcentajes entre ambos trabajos podrían explicarse por el estado de salud de los murciélagos analizados y / o el uso de diferentes tipos de muestras biológicas. El tipo de muestra empleada en el estudio de los murciélagos capturados en Alemania fueron homogeneizados de tejidos de órganos internos en murciélagos muertos o heridos pero en el



caso del estudio húngaro las muestras procedían de EOR murciélagos sanos y muestras de guano en lugares de descanso, por ello la similitud entre el porcentaje de AdV detectado en murciélagos entre este estudio y el realizado con murciélagos capturados en Hungría podría explicarse por la similitud de las muestras empleadas (EOR y H) al ser procedentes de murciélagos sanos.

Es destacable la ausencia de detección de AdV en algunas especies de murciélagos, como *Miniopterus schreibersii*, a pesar de la gran cantidad de individuos de esta especie analizados. Resultados negativos similares se describieron en Alemania y Hungría [141]. La mayoría de los murciélagos en los que se detectó AdV se encuentran dentro de la diversa familia de murciélagos Vespertilionidae y particularmente dentro de la tribu *Pipistrellini* (*Pipistrellus* y *Nyctalus*), mientras que no han sido detectados en la tribu de murciélagos *Plecotinii* (*Barbastella* y *Plecotus*). Dentro de la subfamilia *Myotinae*, se detectaron murciélagos positivos para AdV en varias especies de este grupo. Excepto por el grupo formado por 2 AdV detectados en murciélagos *Myotis emarginatus* (MF540608, MF540609) en el análisis filogenético del fragmento del gen de hexón, el resto de AdV detectados en murciélagos *Myotis* estaban dispersos a lo largo de los árboles filogenéticos sin formar ningún grupo monofilético. Curiosamente, no se detectaron AdV en algunos murciélagos *Myotis*, por ejemplo *M. daubentonii*, a pesar de que dicha especie estaba bien representada en el análisis (n = 60 y n = 41 para EOR y H, respectivamente).

Los estudios anteriores se centraron principalmente en el análisis de guano y tejidos internos [136,137,140,141,144]. El análisis de EOR para la detección de AdV es un aspecto novedoso de esta tesis que ha permitido la detección de AdV en el tracto respiratorio superior de los murciélagos y prueba una posible ruta de transmisión fecal-oral de estos virus. En 2 murciélagos de la especie *P. pygmaeus* (Ppy15 160725 y Ppy14 60725) fue detectado el mismo AdV tanto en muestra de EOR como H apoyando la existencia de esta posible transmisión oral-fecal. Una razón importante para el estudio de EOR en murciélagos es el hecho de que muchos serotipos de AdV humano no tienen un receptor celular específico bien identificado y, dado que se replican mal en animales [154], la comprensión de los factores que definen el tropismo y la transmisión durante una infección natural puede aumentar el conocimiento de las infecciones por AdV. En particular, la existencia de una posible ruta de transmisión fecal-oral de AdV es un tema interesante para explorar considerando a los murciélagos como vectores de enfermedades infecciosas emergentes y reemergentes.

Trabajos previos han publicado potenciales nuevos mastadenovirus de murciélago basándose principalmente en el análisis filogenético de un fragmento corto e informativo del gen de la ADN polimerasa ADN dependiente [137–139,141]. Este gen es un gen bien conservado implicado en la transcripción viral [251], a pesar de su uso habitual en el análisis filogenético de AdV humanos y animales la resolución del análisis filogenético basado en este fragmento es limitada (menos de 100 aminoácidos). La PCR presentada en este estudio amplificó ~ 450 pb, ofreciendo la posibilidad de aumentar la resolución del análisis filogenético. Sin embargo, con el objetivo de comparar nuestras secuencias con las previamente publicadas del centro de Europa [141] y las secuencias de referencia disponibles en la base de datos *GenBank*, la longitud de la secuencia analizada se redujo a 277 nt. Actualmente, el ICTV ha aceptado 2 propuestas de especies de AdV de murciélago: el mastadenovirus A de murciélago [137] y el mastadenovirus B de murciélago [140]. De acuerdo con los criterios taxonómicos [135] y en base al análisis de distancia entre pares los mastadenovirus de murciélago identificados en este objetivo representan especies potencialmente nuevas del género *Mastadenovirus* y muy divergentes con respecto los mastadenovirus de murciélago de referencia del ICTV con la excepción de 3 AdVs detectados en murciélagos de la especie *P. pygmaeus* (Ppy15 160725, Ppy21 160725, Ppy14 160725). Además, un AdV detectado en un murciélago *P. kuhlii* (Pku2\_160622) y otro detectado en un murciélago *N. lasiopterus* (Nlas\_K01317) se asociaron con el mastadenovirus B de murciélago, aunque la obtención de más del 15% de diferencia en la secuencia de aa del análisis de distancia entre pares sugiere que también son potenciales nuevas especies. Es destacable la clasificación en dos grupos diferentes bien apoyados estadísticamente de los AdV de murciélagos obtenidos de las especies *P. kuhlii* y *P. pygmaeus* mostrando la especificidad de estos virus con respecto al huésped incluso a nivel de especie.

En esta tesis la identificación de nuevos AdVs de murciélagos se ve respaldada por los resultados obtenidos utilizando el fragmento del gen de hexón, una proteína más variable [138,139,143,146] que contiene siete regiones hipervariables identificadas como epítomos virales [252]. La variación de los ácidos nucleicos de esta proteína define los diferentes serotipos humanos [253]. La PCR convencional genérica Hexón para AdV (Métodos Apartado 4.13.3) fue diseñada para la amplificar el fragmento en la región hipervariable 7 y el análisis de las secuencias obtenidas estuvo en concordancia con el genotipo y el serotipo en AdV humanos [254].

Las relaciones evolutivas basadas en los dos genes parciales analizados se presentan por separado, ya que proporcionan información diferente según sus diferentes tasas de mutación. Ambos genes coinciden en la estructura principal de sus topologías y grupos de árboles y ambos poseen respaldo estadístico suficiente para un nuevo grupo de mastadenovirus de murciélagos ibéricos que agrupa y distingue entre las familias Vespertilionidae y Rhinolophidae en las filogenias. La mayoría de los AdV disponibles en la base de datos *GenBank* se agruparon dentro de los tres grupos monofiléticos correspondientes al género de su hospedador: *Pipistrellus*, *Nyctalus* y *Rhinolophus*. Esta relación está respaldada además por el análisis filogenético del fragmento del gen de la ADN polimerasa en el cual el AdV detectado en un murciélago *N. leisleri* (Nle 00954) agrupa con un AdV de murciélago detectado en otro murciélago *N. leisleri* muestreado en Hungría [141]. En nuestro muestreo, las relaciones basales entre los principales murciélagos hospedadores fueron más difíciles de asociar con los virus que hospedan debido a la falta de representación de grupos importantes de murciélagos como *Scotophilinii*, *Nycticeinii* y *Plecotinii* dentro de la familia Vespertilionidae. La relación huésped-patógeno en murciélagos se observó claramente en un trabajo previo del laboratorio de Virus Respiratorios con los virus herpes [129] pero aun así los AdV podrían representar otro ejemplo de evolución paralela del virus ADN y sus huéspedes. El análisis filogenético del fragmento del gen de hexón no mostró diferencias en las secuencias obtenidas de AdV de murciélagos capturados en el sur y el norte de España, como se muestra en el grupo que conforman los AdV detectados en *P. pygmaeus* (Ppy 070616) cuya muestra fue recolectada en Lugo y el AdV detectado en un murciélago *P. pygmaeus* (Ppy15 160725) recogido en Sevilla.

Aunque la mayoría de los AdV detectados fueron agrupados por hospedador algunas excepciones son claramente notables. En el análisis filogenético del fragmento del gen de hexón, el AdV detectado en un *M. myotis* (Mmy8 080623) se relacionó con el grupo compuesto por 2 AdVs detectados en especies diferentes de murciélagos *Rhinolophus*. Es conocido que muchas colonias de murciélagos *Myotis* comparten refugio con varias especies del género *Rhinolophus* pudiendo ser este el origen de la transmisión interespecífica entre estas 2 especies de murciélagos. En segundo lugar, el AdV detectado en un *H. savii* (Hsa6 070704) que fue agrupado con los AdVs del grupo *Rhinolophus* en el análisis filogenético del hexón y que en el análisis del gen de la ADN polimerasa formaba parte de un grupo de AdV específico detectados en 3 murciélagos de la especie *H. savii* diferentes. En este segundo ejemplo una transmisión natural parece menos

probable ya que las dos especies tienen una historia de vida muy diferente y apenas comparten ningún requisito ecológico. Sin embargo, la descripción de virus recombinantes es un fenómeno común en AdV humano [255] y podría explicar los resultados obtenidos. La confirmación de una esta recombinación en este AdV de murciélago requiere del análisis de la secuencia genómica completa. Una tercera excepción mostró el AdV detectado en un *M. emarginatus* (Mem15 080703) agrupado junto con el mastadenovirus G de murciélago detectado en la especie *Corynorhinus rafinesquii* y otros 2 AdV detectados en murciélagos *Myotis* capturados en Hungría. El *C. rafinesquii* es un murciélago vespertiliónido, cuya distribución está restringida al sudeste de América del Norte y México [146,256]. La conexión entre estos virus es una incógnita dado que sus anfitriones están distantes geográficamente y evolutivamente, pero podría estar relacionado este fenómeno con una reciente colonización de América del Norte por *Myotis* paleárticos [257].

En conclusión, basado en el análisis de dos regiones diferentes del genoma utilizadas para estudiar dos tipos diferentes de muestras, el presente estudio contribuye con miembros potencialmente nuevos del género Mastadenovirus, distintos de las especies de referencia previamente descritas mastadenovirus A y B de murciélagos [137,139]. Los nuevos grupos de AdV se detectaron en murciélagos capturados en una amplia región geográfica y generan datos que respaldan que la diversidad de mastadenovirus de murciélago está asociada con el hospedador y la distribución éste.

La obtención de genoma completo de virus ADN de gran tamaño sigue siendo uno de los restos de la secuenciación de nueva generación. En el caso de AdV trabajos previos empleando métodos NGS han sido publicados aunque ninguno de ellos fue realizado en muestras respiratorias [148,258] ni de forma directa en estas muestras respiratorias [147].

El enriquecimiento mediante captura por sondas se ha considerado mucho más efectivo que el enriquecimiento por 3 métodos en la secuenciación del genoma completo de virus [151]. En el desarrollo de este objetivo comparamos los resultados obtenidos tras el análisis bioinformático de las secuencias obtenidas tras el enriquecimiento por 3 métodos con la preparación de librerías con el kit *Nextera XT* (Materiales. Apartado 3.5.2) y las secuencias obtenidas tras la preparación de librerías con el kit *Kapa Hyper Plus* y el enriquecimiento mediante captura por sondas con el kit *HyperCap Target Enrichment* (Materiales. Apartado 3.5.2).

Los resultados obtenidos reflejaron que el número de fragmentos virales detectados tras la identificación taxonómica fue similar en ambos procedimientos (3,2 millones) pero la distribución de estos fragmentos fue diferente según el método empleado (Figuras 43 y 44). En el caso del enriquecimiento por 3 métodos el porcentaje de número de fragmentos de secuencias de virus fue de un 22 % frente al 17% obtenidos con el método de captura por sondas, a pesar de esto el número de fragmentos obtenidos correspondiente a AdV, 12,5% del total frente al 30%, fue menor en comparación con los resultados obtenidos tras el enriquecimiento mediante captura por sondas. La presencia de mayor cantidad de bacteriófagos y virus CRESS-DNA incrementó el porcentaje de fragmentos de secuencia de virus en el caso del enriquecimiento por 3 métodos, esto puede ser debido a que el filtrado puede influir en la detección de virus de gran tamaño favoreciendo el aumento de proporción de virus de pequeño tamaño [192].

Durante el desarrollo de este objetivo fue corroborado que los mejores resultados se obtuvieron tras el enriquecimiento por captura. El empleo del enriquecimiento mediante captura por sondas permitió la obtención de mayores tamaños de la secuencia de AdV (Tabla 39. Página 202). En el caso del AdV “Ppy15 181106” fueron obtenidos 16884 de los 31616 que tiene el BtAdV-2 (*GenBank*: JN252129). El fragmento obtenido fue depositado en la base de datos *GenBank* quedando identificado como MN490088.

Debido a la dificultad de la WGS de AdV no hay muchas secuencias completas de adenovirus de murciélago disponibles en plataformas como *GenBank* (Materiales. Apartado 3.8). La mayoría de las secuencias disponibles proceden del análisis de cultivos celulares infectados con virus de murciélagos que procedían de órganos internos de animales enfermos [133,146]. Por el contrario, las muestras empleadas en esta tesis provienen de muestras respiratorias de murciélagos sanos por lo que la detección y secuenciación del virus es más complicada.

La baja proporción de virus en las muestras respiratorias puede ser una de las razones por las que el cultivo celular no ha sido exitoso.

Las causas de la falta de secuencia completa a pesar de un enriquecimiento eficaz pueden ser multifactoriales: puede deberse al tipo de muestra, al estado de la muestra o a la técnica empleada. Dar solución al cualquiera de las cuestiones anteriores sigue siendo un rato al que dar respuesta.

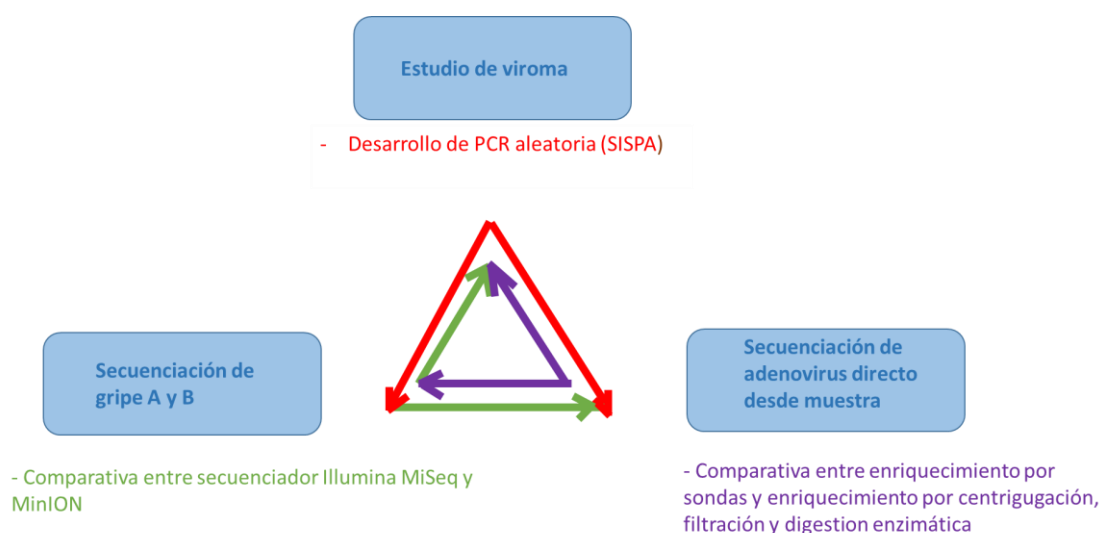
# EPÍLOGO



## EPÍLOGO

El desarrollo de los métodos de NGS que se han ajustado y estandarizado en esta tesis doctoral han demostrado ser métodos eficaces y sensibles. Empleados en el estudio del viroma del tracto respiratorio y en la posterior caracterización de los virus presentes en dicho viroma resultaron extremadamente útiles, incluso cuando se realizaron sobre las muestras clínicas directamente, estos métodos aportaron gran cantidad de información genómica dado que muchos de los genes virales fueron secuenciados de manera eficaz. A pesar de tener una gran sensibilidad presentaron en determinados casos dificultades para obtener el genoma completo de algunos de los virus respiratorios.

Los diferentes objetivos de esta tesis no son compartimentos estancos independientes. La estandarización, la implementación y el análisis de resultados obtenidos plantea un aumento del conocimiento de la caracterización del viroma del tracto respiratorio que es el hilo conector que nos lleva a desarrollar nuevos protocolos de trabajo y adaptar los que existen para poder comprender la ecología de la infección respiratoria abarcando diferentes virus respiratorios (Figura 47).



**Figura 47:** Relación entre las tecnologías estandarizadas para cada estudio definido en los objetivos de trabajo de esta tesis y su influencia. Las líneas rojas representan la influencia que ha tenido el desarrollo del método de amplificación independiente de secuencia en la secuenciación de gripe y de adenovirus. Las líneas verdes representan la influencia que la comparación entre secuenciadores para muestras de gripe ha tenido en la secuenciación de las muestras de viroma y de adenovirus. Las líneas moradas representan la influencia de los diferentes métodos de enriquecimiento en muestras de murciélagos y muestras respiratorias humanas empleadas en el estudio de viroma y de secuenciación de gripe.



En la estandarización del flujo de trabajo para el estudio de viroma se realizó el estudio sobre el efecto del enriquecimiento viral en la secuenciación NGS de la muestra respiratoria. Los resultados obtenidos mostraron que el empleo de métodos enriquecimiento viral son críticos para la secuenciación de virus respiratorios. En las secuencias obtenidas de las muestras no enriquecidas o no se pudo detectar virus respiratorios o el tamaño de secuencia obtenido era menor que el secuenciado empleando el enriquecimiento viral sobre las muestras (Tabla 16. Página 118).

Los métodos de enriquecimiento viral previamente publicados son variados y realizados sobre diferente tipo de muestras. Se han propuestos métodos basados en la separación del citosol del resto de la célula y la recuperación de la fracción citosólica donde se encuentran los virus encapsulados para el estudio de biopsias [180], métodos basados en tratamiento con ditiotreitol, filtración y digestión enzimáticas en heces [190] y en el caso del estudio de muestras clínica la filtración, digestión por nucleasas y la amplificación SISPA ADN y ARN en reacciones separadas fueron propuestos como los métodos que aportan mejor resultado en el estudio de muestras clínicas [191].

Aunque trabajos previos destacan la relevancia del empleo de enriquecimiento viral es importante que la elección del método sea acorde con el tipo de muestra de estudio [9] y con las características morfológicas de los virus respiratorios, ya que el filtrado con filtro de 0.22  $\mu\text{m}$  puede eliminar virus grandes [192] o el empleo de ADNasa I puede eliminar virus ADN de cadena simple [259].

El empleo de métodos de enriquecimiento mediante captura por sondas presenta ventajas frente a otros métodos debido a la eliminación de técnicas que pueden afectar a la representación de la diversidad viral de la muestra debido a la variedad de tamaños y morfologías de los virus presentes en el tracto respiratorio.

Los resultados obtenidos comparando el método de enriquecimiento por centrifugación, filtrado y digestión con nucleasas con el enriquecimiento por captura por sondas en ANF de pacientes pediátricos con RV mostraron que aunque la detección de virus por ambos métodos es similar el enriquecimiento por captura con sondas permite la obtención de genomas completos o al menos de mayor tamaño que los obtenidos empleando el método de enriquecimiento por 3 métodos.

El empleo de ambos métodos son útiles según el objetivo del estudio, el enriquecimiento mediante captura por sondas presenta ventajas a la hora de la obtención de mayores tamaños de secuencias virales pero no puede detectar secuencias con una homología menor del 75% con las sondas empleadas [151] y presenta menos sensibilidad en la detección de los virus CRESS-DNA que el método de enriquecimiento por 3 métodos.

El empleo del método de amplificación aleatoria SISPA presenta algunos problemas como la falta de homogeneidad en la cobertura de las secuencias sobrerrepresentando algunas regiones, aun así el impacto de este sesgo en la diversidad viral es menor [192]. Lewandowska y col. [191] propusieron en su trabajo que el empleo de reacciones diferentes de SISPA ADN y ARN eran el mejor método posible para la detección de todos los virus presentes en la muestra. El protocolo de trabajo estandarizado en esta tesis realiza las reacciones por separado con el fin de reducir el número de librerías a realizar. La mezcla equimolar de ambos productos de reacción no produce una pérdida de representatividad de la diversidad viral detectada en las muestras analizadas (Tabla 17. Página 120).

La estandarización de los protocolos para el estudio de viroma respiratorio permitió el estudio de muestra respiratoria de pacientes que presentaban síntomas de infección respiratoria pero no les fue diagnosticado ningún virus en las PCRs tiempo real de diagnóstico. Con el fin de detectar los virus presentes en esas muestras se realizó el enriquecimiento por 3 métodos, amplificación con SISPA y preparación de librerías con el kit *Nextera XT*. De los 57 pacientes sin patógeno descrito, en 35 de ellos fue identificado el patógeno viral causante de la infección (Figura 18. Página 128). El mejor resultado de los métodos NGS fue debido al bajo rendimiento de los métodos moleculares y no al hallazgo de nuevos patógenos. Taboada y col. [213] analizaron LNF de 25 pacientes ingresados en el hospital con neumonía por NGS, este trabajo detectó patógenos conocidos en 22 de las 25 muestras de los pacientes analizados presentando resultados similares a los reflejados en este trabajo.

El análisis del grupo control por este método permitió determinar los virus que están presentes en el viroma respiratorio de pacientes sanos y evaluar la relevancia clínica de los virus identificados en las muestras de pacientes con infección respiratoria (Tabla 20. Página 132).

Debido al desarrollo de métodos para el estudio de viroma obtuvimos datos interesantes que aportan información novedosa sobre un grupo muy complejo por su diversidad biológica como son los rinovirus. El estudio de muestras de seguimiento procedentes de pacientes pediátricos con infección respiratoria aguda ha permitido caracterizar el grupo de rinovirus mayoritario en los pacientes analizados e incluso se ha podido determinar genómicamente los tipos de virus presentes en la muestra a lo largo del tiempo. El empleo del método de enriquecimiento mediante captura por sondas permitió el análisis de la secuencia de nucleótidos del gen VP1 de RVA que reveló la existencia de procesos infecciosos independientes con diferente tipo de RVA en muestras con un mes de diferencia (Tabla 22. Página 135). En el caso de RVB tras un mes se seguía detectando el mismo virus. Los RV se consideran virus de larga persistencia determinándose la horquilla de detección entre 5 y 6 semanas [221]. Waman y col. [224] propuso que la existencia de subpoblaciones en RVA podría ser la explicación de la variabilidad de estos virus frente a la homogeneidad de RVB.

El aumento del conocimiento sobre virus circulares va asociado al auge de estas tecnologías de secuenciación. El análisis de las muestras pediátricas mostró la presencia mayoritaria del virus TTV en muestras de pacientes con infección respiratoria positivos para RV frente a individuos control sanos (Figura 20. Página 131). La influencia de estos virus en infección respiratoria no está clara aunque han sido asociados a sibilancias y asma [231]. Los resultados mostrados son un punto de partida, estudios futuros de cohorte con un número mayor de pacientes son necesarios para demostrar una posible asociación entre estos virus.

Los métodos empleados son un apoyo fundamental para la caracterización genómica de los virus caracterizados en los Laboratorios de Referencia. Cuando las técnicas de detección de los virus respiratorios que tradicionalmente se utilizan en dichos laboratorios no permiten la filiación de las infecciones, la posibilidad de realizar una investigación mediante estos métodos para poder conocer la existencia de otros posibles virus en el tracto respiratorio es una gran ventaja. El análisis de las muestras respiratorias y no respiratorias de una paciente ingresada en UCI con diagnóstico negativo a todos los virus respiratorios conocidos, permitió el hallazgo de un nuevo virus (Tabla 24. Página 151). La secuenciación de su genoma mediante métodos RCA y PCRs específicas de amplificación de genoma completo permitió la caracterización genómica y su ubicación dentro de los virus CRESS-DNA. A pesar de que presenta diferencias en el nonúmero conservado y en algunas estructuras conservadas, el virus forma parte de la

nueva familia *Redondoviridae* [99]. Trabajos sobre la infección de estos virus y su papel en la IRA son el futuro de la caracterización de estos virus HRCiV.

El desarrollo de métodos de secuenciación masiva sencillos, rápidos y fácilmente implementables en los Laboratorios de Referencia de Gripe han permitido la secuenciación simultánea de los ocho segmentos del genoma de los virus gripales. Este hecho es de enorme importancia para conocer rápidamente las características genómicas de los virus gripales estacionales, las características de los virus de la gripe responsables de pandemias en el futuro o las adaptaciones de los virus de la gripe procedentes de otros animales en el ser humano. La estandarización de la WGS del virus de la gripe permitió la detección de una coinfección de gripe A/H3N2 y gripe B linaje Yamagata en dos hermanos pediátricos con infección respiratoria que no había sido detectada por los métodos moleculares tradicionalmente empleados en la caracterización de los virus de la gripe.

El estudio de los AdVs presentes en murciélagos y las relaciones filogenéticas existentes analizando dos fragmentos de dos genes diferentes ubicó a los AdVs detectados dentro del grupo de los Mastadenovirus. Estos virus son distintos de los mastadenovirus de murciélago de referencia por lo que podrían ser potenciales nuevas especies. Los grupos filogenéticos formados respaldan que la diversidad de mastadenovirus de murciélagos está asociada con el hospedador del que se obtuvieron las muestras analizadas (Figuras 41 y 42).

El empleo de muestras respiratorias de murciélago muestra que las técnicas desarrolladas son específicas para muestra respiratoria de cualquier tipo de mamífero.

Utilizando los potenciales nuevos mastadenovirus detectados en murciélagos ibéricos se realizaron pruebas de enriquecimiento viral para favorecer la secuenciación del genoma completo virus ADN de gran tamaño como son los AdVs.

La obtención de secuencia completa de AdVS por un lado es imprescindible para su propuesta taxonómica como nuevo adenovirus al ICTV y por otro lado permite poner a punto un método de secuenciación directa de muestra sin necesidad de cultivo previo para virus ADN de gran tamaño. El mayor tamaño de secuencia obtenido empleando el método de enriquecimiento mediante captura por sondas fue de 16884 nt, aproximadamente la mitad del genoma de estos virus. Durante el desarrollo de este objetivo la secuenciación NGS se realizó desde muestra respiratoria directa. Trabajos previos que secuenciaron el genoma completo de

AdVpartieron del aislamiento viral en cultivo celular de muestras de murciélagos enfermos [142,146]. En nuestro caso no fue posible el aislamiento viral de estos virus de murciélago debido a la baja carga viral de las muestras respiratorias que fueron tomadas en murciélagos aparentemente sanos.

Este trabajo presenta algunas respuestas a retos metodológicos para mejorar el conocimiento de los virus del tracto respiratorio y su aplicación a pacientes críticos, gracias al estudio y desarrollo de estos métodos describimos técnicas aplicables en cualquier tipo de estudio sobre muestra respiratoria humana o animal. Objetivo a objetivo se ha consolidado la metodología y obtenido respuestas a algunas preguntas planteadas al inicio de la tesis pero aún quedan incógnitas por resolver siendo esta tesis el primer escalón de un camino lleno de futuro

# CONCLUSIONES



- 1) La estandarización del flujo de trabajo empleando el enriquecimiento por 3 métodos (centrifugación, filtración y digestión con nucleasas) permitió la detección de virus nuevos y conocidos en el estudio de viroma respiratorio. El ajuste del flujo de trabajo empleando el enriquecimiento mediante captura por sondas permitió la secuenciación de genomas completos y grandes tamaños de secuencia de virus conocidos.
- 2) La secuenciación por NGS de 57 muestras respiratorias de pacientes con infección y diagnóstico negativo, reveló la presencia de virus respiratorios comúnmente conocidos en 35 de las muestras analizadas. Sin embargo, en 22 de ellas no se identificó ningún virus conocido.
- 3) El análisis de 61 niños control sanos permitió la caracterización del viroma respiratorio en ausencia de infección respiratoria siendo detectados virus de las familias *Papillomaviridae*, *Anelloviridae*, *Picornaviridae*, *Herpesviridae* y *Reoviridae*.
- 4) El estudio de 35 muestras de seguimiento de 11 pacientes con infección respiratoria diagnosticada por detección de rinovirus, reveló que el 64% de los virus detectados fue el RV grupo A. El análisis de la secuencia de nucleótidos del gen VP1 mostró que los RVA analizados un mes después de la primera toma de muestra, son virus diferentes asociados a episodios de infección independientes.
- 5) Los resultados obtenidos del estudio de viroma de los 11 pacientes infectados por RV, revelaron que TTV está presente en 5 de los 11 pacientes colocándolo como el segundo grupo de virus mayormente detectado. El historial clínico de 4 de estos 5 pacientes reveló hospitalización por IRA. Adicionalmente, en el análisis de los 21 individuos control sanos de seguimiento no fue detectado ningún TTV.
- 6) El uso de métodos NGS para el estudio y caracterización del viroma respiratorio permitió el descubrimiento de un nuevo virus circular presente en el tracto respiratorio de pacientes ingresados en UCI por neumonía grave denominado *Human Respiratory Circular Virus* (HRCiV).
- 7) La obtención del genoma completo del virus HRCiV tras la amplificación por círculo rodante y el análisis del genoma completo, demuestran la circularidad del virus y la presencia de estructuras conservadas, noámero conservado o los motivos conservados de proteínas relacionadas con la replicación por círculo rodante, identificaron definitivamente a este virus como CRESS-DNA.



- 8) Virus similares a HRCiV fueron detectados en 28 pacientes adultos, de los cuales 5 presentaban diagnóstico de neumonía grave y resultaron negativos para el resto de los virus respiratorios. La presencia del virus HRCiV en pacientes pediátricos es muy baja. La ausencia en controles negativos y pacientes pediátricos descarta "*a priori*" el papel como virus comensal del tracto respiratorio.
- 9) Analizando las secuencias de los 8 segmentos genómicos de los virus de la gripe A y B de manera simultánea y directamente detectados en la muestra clínica, permite la caracterización de variantes minoritarias que presentan un altísimo interés clínico en la descripción de los virus asociados al desarrollo de un proceso gripal.
- 10) Los métodos de secuenciación de genoma completo permitieron la detección y caracterización de una coinfección por gripe A/H3N2 y gripe B.
- 11) El empleo del método de enriquecimiento mediante captura por sondas permitió el análisis de fragmentos de mayor tamaño de secuencia de nuevas especies del género Mastadenovirus que empleando otros métodos de enriquecimiento viral.
- 12) El empleo de los métodos desarrollados en esta tesis permite la caracterización del viroma respiratorio humano y animal, lo que supone un avance tecnológico que mejorará el conocimiento de la infección respiratoria viral por su directa implementación.

# BIBLIOGRAFÍA



1. Chiu, C.Y.; Miller, S.A. Clinical metagenomics. *Nature Reviews Genetics* **2019**.
2. Sanger, F.; Coulson, A.R. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *Journal of Molecular Biology* **1975**, *94*, 441–448.
3. Miller, M.B.; Tang, Y.-W. Basic Concepts of Microarrays and Potential Applications in Clinical Microbiology. *Clin Microbiol Rev* **2009**, *22*, 611–633.
4. Rota, P.A.; Oberste, M.S.; Monroe, S.S.; Nix, W.A.; Campagnoli, R.; Icenogle, J.P.; Peñaranda, S.; Bankamp, B.; Maher, K.; Chen, M.; et al. Characterization of a Novel Coronavirus Associated with Severe Acute Respiratory Syndrome. *Science* **2003**, *300*, 1394–1399.
5. Malla, M.A.; Dubey, A.; Kumar, A.; Yadav, S.; Hashem, A.; Abd\_Allah, E.F. Exploring the Human Microbiome: The Potential Future Role of Next-Generation Sequencing in Disease Diagnosis and Treatment. *Frontiers in Immunology* **2019**, *9*.
6. Pérez-Breña P.; Aguilar Ruiz J. La importancia de los sistemas de vigilancia gripal. *Revisiones sistemáticas de la Sociedad Española de Enfermedades Infecciosas*.
7. Imai, K.; Tamura, K.; Tanigaki, T.; Takizawa, M.; Nakayama, E.; Taniguchi, T.; Okamoto, M.; Nishiyama, Y.; Tarumoto, N.; Mitsutake, K.; et al. Whole Genome Sequencing of Influenza A and B Viruses With the MinION Sequencer in the Clinical Setting: A Pilot Study. *Front Microbiol* **2018**, *9*, 2748.
8. Lu, H.; Giordano, F.; Ning, Z. Oxford Nanopore MinION Sequencing and Genome Assembly. *Genomics, Proteomics & Bioinformatics* **2016**, *14*, 265–279.
9. Hall, R.J.; Wang, J.; Todd, A.K.; Bissielo, A.B.; Yen, S.; Strydom, H.; Moore, N.E.; Ren, X.; Huang, Q.S.; Carter, P.E.; et al. Evaluation of rapid and simple techniques for the enrichment of viruses prior to metagenomic virus discovery. *Journal of Virological Methods* **2014**, *195*, 194–204.
10. Zhang, D.; Lou, X.; Yan, H.; Pan, J.; Mao, H.; Tang, H.; Shu, Y.; Zhao, Y.; Liu, L.; Li, J.; et al. Metagenomic analysis of viral nucleic acid extraction methods in respiratory clinical samples. *BMC Genomics* **2018**, *19*, 773.
11. Reyes, G.R.; Kim, J.P. Sequence-independent, single-primer amplification (SISPA) of complex DNA populations. *Mol. Cell. Probes* **1991**, *5*, 473–481.
12. Palacios, G.; Quan, P.; Jabado, O.J.; Conlan, S.; Hirschberg, D.L.; Liu, Y.; Zhai, J.; Renwick, N.; Hui, J.; Hegyi, H.; et al. Panmicrobial oligonucleotide array for diagnosis of infectious diseases. *Emerging Infect. Dis.* **2007**, *13*, 73–81.
13. Pérez-Sautu, U.; Wiley, M.R.; Iglesias-Caballero, M.; Pozo, F.; Prieto, K.; Chitty, J.A.; García-García, M.L.; Calvo, C.; Casas, I.; Palacios, G. Target-independent high-throughput sequencing methods provide evidence that already known human viral pathogens play a main role in respiratory infections with unexplained etiology. *Emerg Microbes Infect* **2019**, *8*, 1054–1065.
14. Gu L.; Yan, W.; Liu, L.; Wang, S.; Zhang, X.; Lyu, M. Research Progress on Rolling Circle Amplification (RCA)-Based Biomedical Sensing. *Pharmaceuticals (Basel)* **2018**, *11*.
15. Mohsen, M.G.; Kool, E.T. The Discovery of Rolling Circle Amplification and Rolling Circle Transcription. *Acc. Chem. Res.* **2016**, *49*, 2540–2550.
16. Zhou, B.; Donnelly, M.E.; Scholes, D.T.; St. George, K.; Hatta, M.; Kawaoka, Y.; Wentworth, D.E. Single-Reaction Genomic Amplification Accelerates Sequencing and Vaccine Production for Classical and Swine Origin Human Influenza A Viruses. *J Virol* **2009**, *83*, 10309–10313.
17. Zhou, B.; Lin, X.; Wang, W.; Halpin, R.A.; Bera, J.; Stockwell, T.B.; Barr, I.G.; Wentworth, D.E. Universal Influenza B Virus Genomic Amplification Facilitates Sequencing, Diagnostics, and Reverse Genetics. *J Clin Microbiol* **2014**, *52*, 1330–1337.
18. Zhao, J.; Liu, J.; Vemula, S.V.; Lin, C.; Tan, J.; Ragupathy, V.; Wang, X.; Mbondji-wonje, C.; Ye, Z.; Landry, M.L.; et al. Sensitive Detection and Simultaneous Discrimination of Influenza A and B Viruses in Nasopharyngeal Swabs in a Single Assay Using Next-Generation Sequencing-Based Diagnostics. *PLoS One* **2016**, *11*.

19. Oakeson, K.F.; Wagner, J.M.; Mendenhall, M.; Rohrwasser, A.; Atkinson-Dunn, R. Bioinformatic Analyses of Whole-Genome Sequence Data in a Public Health Laboratory. *Emerging Infect. Dis.* **2017**, *23*, 1441–1445.
20. Ali, R.; Blackburn, R.M.; Kozlakidis, Z. Next-Generation Sequencing and Influenza Virus: A Short Review of the Published Implementation Attempts. *HAYATI Journal of Biosciences* **2016**, *23*, 155–159.
21. Shepard, S.S.; Meno, S.; Bahl, J.; Wilson, M.M.; Barnes, J.; Neuhaus, E. Viral deep sequencing needs an adaptive approach: IRMA, the iterative refinement meta-assembler. *BMC Genomics* **2016**, *17*, 708.
22. Goldstein, E.J.; Harvey, W.T.; Wilkie, G.S.; Shepherd, S.J.; MacLean, A.R.; Murcia, P.R.; Gunson, R.N. Integrating patient and whole-genome sequencing data to provide insights into the epidemiology of seasonal influenza A(H3N2) viruses. *Microbial Genomics* **2018**, *4*.
23. Lee, H.K.; Lee, C.K.; Tang, J.W.-T.; Loh, T.P.; Koay, E.S.-C. Contamination-controlled high-throughput whole genome sequencing for influenza A viruses using the MiSeq sequencer. *Scientific Reports* **2016**, *6*, 33318.
24. Yamashita, A.; Sekizuka, T.; Kuroda, M. VirusTAP: Viral Genome-Targeted Assembly Pipeline. *Front Microbiol* **2016**, *7*.
25. Borges, V.; Pinheiro, M.; Pechirra, P.; Guiomar, R.; Gomes, J.P. INSaFLU: an automated open web-based bioinformatics suite “from-reads” for influenza whole-genome-sequencing-based surveillance. *Genome Med* **2018**, *10*.
26. Burd, E.M. Validation of laboratory-developed molecular assays for infectious diseases. *Clin. Microbiol. Rev.* **2010**, *23*, 550–576.
27. Virgin, H.W.; Wherry, E.J.; Ahmed, R. Redefining chronic viral infection. *Cell* **2009**, *138*, 30–50.
28. Duerkop, B.A.; Hooper, L.V. Resident viruses and their interactions with the immune system. *Nat. Immunol.* **2013**, *14*, 654–659.
29. GBD 2015 LRI Collaborators Estimates of the global, regional, and national morbidity, mortality, and aetiologies of lower respiratory tract infections in 195 countries: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet Infect Dis* **2017**, *17*, 1133–1161.
30. WHO World Health Organization. Global Health Estimates 2016: Deaths by Cause, Age, Sex, by Country and by Region, 2000-2016. **2018**.
31. Rudan, I.; O’Brien, K.L.; Nair, H.; Liu, L.; Theodoratou, E.; Qazi, S.; Lukšić, I.; Fischer Walker, C.L.; Black, R.E.; Campbell, H.; et al. Epidemiology and etiology of childhood pneumonia in 2010: estimates of incidence, severe morbidity, mortality, underlying risk factors and causative pathogens for 192 countries. *J Glob Health* **2013**, *3*, 010401.
32. Rey, C.C.; García, M.L.G.; Flecha, I.C.; Breña, P.P. Infecciones respiratorias virales. 16.
33. García-García, M.L.; Calvo, C.; Pérez-Breña, P.; De Cea, J.M.; Acosta, B.; Casas, I. Prevalence and clinical characteristics of human metapneumovirus infections in hospitalized infants in Spain. *Pediatr. Pulmonol.* **2006**, *41*, 863–871.
34. Calvo Rey, C.; García García, M.; Casas Flecha, I.; Martín del Valle, F.; Centeno Jiménez, M.; Pérez-Breña, P. Infecciones por virus de la gripe en menores de dos años. *An Pediatr (Barc)* **2005**, *63*, 22–28.
35. García-García, M.L.; Calvo, C.; Rey, C.; Díaz, B.; Molinero, M.D.M.; Pozo, F.; Casas, I. Human metapneumovirus infections in hospitalized children and comparison with other respiratory viruses. 2005-2014 prospective study. *PLoS ONE* **2017**, *12*, e0173504.
36. García-García, M.L.; Calvo, C.; Falcón, A.; Pozo, F.; Pérez-Breña, P.; De Cea, J.M.; Casas, I. Role of emerging respiratory viruses in children with severe acute wheezing. *Pediatr. Pulmonol.* **2010**, *45*, 585–591.
37. Regamey, N.; Kaiser, L.; Roiha, H.L.; Deffernez, C.; Kuehni, C.E.; Latzin, P.; Aebi, C.; Frey, U.; Swiss Paediatric Respiratory Research Group Viral etiology of acute respiratory infections with cough in infancy: a community-based birth cohort study. *Pediatr. Infect. Dis. J.* **2008**, *27*, 100–105.

38. Kusel, M.M.H.; de Klerk, N.H.; Holt, P.G.; Keadze, T.; Johnston, S.L.; Sly, P.D. Role of respiratory viruses in acute upper and lower respiratory tract illness in the first year of life: a birth cohort study. *Pediatr. Infect. Dis. J.* **2006**, *25*, 680–686.
39. Pozo, F.; García-García, M.L.; Calvo, C.; Cuesta, I.; Pérez-Breña, P.; Casas, I. High incidence of human bocavirus infection in children in Spain. *J. Clin. Virol.* **2007**, *40*, 224–228.
40. Calvo, C.; García-García, M.L.; Blanco, C.; Vázquez, M.C.; Frías, M.E.; Pérez-Breña, P.; Casas, I. Multiple simultaneous viral infections in infants with acute respiratory tract infections in Spain. *J. Clin. Virol.* **2008**, *42*, 268–272.
41. Hamparian, V.V.; Colonno, R.J.; Cooney, M.K.; Dick, E.C.; Gwaltney, J.M.; Hughes, J.H.; Jordan, W.S.; Kapikian, A.Z.; Mogabgab, W.J.; Monto, A. A collaborative report: rhinoviruses--extension of the numbering system from 89 to 100. *Virology* **1987**, *159*, 191–192.
42. Bochkov, Y.A.; Gern, J.E. Clinical and molecular features of human rhinovirus C. *Microbes Infect.* **2012**, *14*, 485–494.
43. Savolainen, C.; Blomqvist, S.; Mulders, M.N.; Hovi, T. Genetic clustering of all 102 human rhinovirus prototype strains: serotype 87 is close to human enterovirus 70. *J. Gen. Virol.* **2002**, *83*, 333–340.
44. Savolainen, C.; Mulders, M.N.; Hovi, T. Phylogenetic analysis of rhinovirus isolates collected during successive epidemic seasons. *Virus Res.* **2002**, *85*, 41–46.
45. Dominguez, S.R.; Briese, T.; Palacios, G.; Hui, J.; Villari, J.; Kapoor, V.; Tokarz, R.; Glodé, M.P.; Anderson, M.S.; Robinson, C.C.; et al. Multiplex MassTag-PCR for respiratory pathogens in pediatric nasopharyngeal washes negative by conventional diagnostic testing shows a high prevalence of viruses belonging to a newly recognized rhinovirus clade. *J. Clin. Virol.* **2008**, *43*, 219–222.
46. Song, D.J. Rhinovirus and childhood asthma: an update. *Korean J Pediatr* **2016**, *59*, 432–439.
47. Lee, W.-M.; Lemanske, R.F.; Evans, M.D.; Vang, F.; Pappas, T.; Gangnon, R.; Jackson, D.J.; Gern, J.E. Human Rhinovirus Species and Season of Infection Determine Illness Severity. *Am J Respir Crit Care Med* **2012**, *186*, 886–891.
48. Bizzintino, J.; Lee, W.-M.; Laing, I.A.; Vang, F.; Pappas, T.; Zhang, G.; Martin, A.C.; Khoo, S.-K.; Cox, D.W.; Geelhoed, G.C.; et al. Association between human rhinovirus C and severity of acute asthma in children. *Eur. Respir. J.* **2011**, *37*, 1037–1042.
49. Foxman, E.F.; Iwasaki, A. Genome-virome interactions: examining the role of common viral infections in complex disease. *Nat. Rev. Microbiol.* **2011**, *9*, 254–264.
50. Lemanske, R.F.; Jackson, D.J.; Gangnon, R.E.; Evans, M.D.; Li, Z.; Shult, P.A.; Kirk, C.J.; Reisdorf, E.; Roberg, K.A.; Anderson, E.L.; et al. Rhinovirus illnesses during infancy predict subsequent childhood wheezing. *J. Allergy Clin. Immunol.* **2005**, *116*, 571–577.
51. Rubner, F.J.; Jackson, D.J.; Evans, M.D.; Gangnon, R.E.; Tisler, C.J.; Pappas, T.E.; Gern, J.E.; Lemanske, R.F. Early life rhinovirus wheezing, allergic sensitization, and asthma risk at adolescence. *J. Allergy Clin. Immunol.* **2017**, *139*, 501–507.
52. Nishizawa, T.; Okamoto, H.; Konishi, K.; Yoshizawa, H.; Miyakawa, Y.; Mayumi, M. A novel DNA virus (TTV) associated with elevated transaminase levels in posttransfusion hepatitis of unknown etiology. *Biochem. Biophys. Res. Commun.* **1997**, *241*, 92–97.
53. De Vlaminck, I.; Khush, K.K.; Strehl, C.; Kohli, B.; Luikart, H.; Neff, N.F.; Okamoto, J.; Snyder, T.M.; Cornfield, D.N.; Nicolls, M.R.; et al. Temporal response of the human virome to immunosuppression and antiviral therapy. *Cell* **2013**, *155*, 1178–1187.
54. Maggi, F.; Pifferi, M.; Tempestini, E.; Fornai, C.; Lanini, L.; Andreoli, E.; Vatteroni, M.; Presciuttini, S.; Pietrobelli, A.; Boner, A.; et al. TT virus loads and lymphocyte subpopulations in children with acute respiratory diseases. *J. Virol.* **2003**, *77*, 9081–9083.
55. Pifferi, M.; Maggi, F.; Andreoli, E.; Lanini, L.; Marco, E.D.; Fornai, C.; Vatteroni, M.L.; Pistello, M.; Ragazzo, V.; Macchia, P.; et al. Associations between nasal torquetenovirus load and spirometric indices in children with asthma. *J. Infect. Dis.* **2005**, *192*, 1141–1148.

56. Prasetyo, A.A.; Desyardi, M.N.; Tanamas, J.; Suradi, null; Reviono, null; Harsini, null; Kageyama, S.; Chikumi, H.; Shimizu, E. Respiratory viruses and torque teno virus in adults with acute respiratory infections. *Intervirology* **2015**, *58*, 57–68.
57. Edwards, R.A.; Rohwer, F. Viral metagenomics. *Nat. Rev. Microbiol.* **2005**, *3*, 504–510.
58. Angly, F.E.; Felts, B.; Breitbart, M.; Salamon, P.; Edwards, R.A.; Carlson, C.; Chan, A.M.; Haynes, M.; Kelley, S.; Liu, H.; et al. The marine viromes of four oceanic regions. *PLoS Biol.* **2006**, *4*, e368.
59. Delwart, E.L. Viral metagenomics. *Rev. Med. Virol.* **2007**, *17*, 115–131.
60. Finkbeiner, S.R.; Allred, A.F.; Tarr, P.I.; Klein, E.J.; Kirkwood, C.D.; Wang, D. Metagenomic analysis of human diarrhea: viral detection and discovery. *PLoS Pathog.* **2008**, *4*, e1000011.
61. Phan, T.G.; Kapusinszky, B.; Wang, C.; Rose, R.K.; Lipton, H.L.; Delwart, E.L. The fecal viral flora of wild rodents. *PLoS Pathog.* **2011**, *7*, e1002218.
62. Rosario, K.; Nilsson, C.; Lim, Y.W.; Ruan, Y.; Breitbart, M. Metagenomic analysis of viruses in reclaimed water. *Environ. Microbiol.* **2009**, *11*, 2806–2820.
63. Shan, T.; Li, L.; Simmonds, P.; Wang, C.; Moeser, A.; Delwart, E. The fecal virome of pigs on a high-density farm. *J. Virol.* **2011**, *85*, 11697–11708.
64. Biagini P; Bendinelli M; Hino S; Kakkola L; Mankertz A; Niel C; Okamoto H; Raidal S; Teo CG; Todd D Family Anelloviridae. In *Virus taxonomy: classification and nomenclature of viruses: ninth report of the International Committee on Taxonomy Viruses.*; Academic Press, London, 2012; pp. 331–341.
65. Lefevre, P.; Lett, J.-M.; Varsani, A.; Martin, D.P. Widely conserved recombination patterns among single-stranded DNA viruses. *J. Virol.* **2009**, *83*, 2697–2707.
66. Martin, D.P.; Biagini, P.; Lefevre, P.; Golden, M.; Roumagnac, P.; Varsani, A. Recombination in eukaryotic single stranded DNA viruses. *Viruses* **2011**, *3*, 1699–1738.
67. Duffy, S.; Shackelton, L.A.; Holmes, E.C. Rates of evolutionary change in viruses: patterns and determinants. *Nat. Rev. Genet.* **2008**, *9*, 267–276.
68. Firth, C.; Charleston, M.A.; Duffy, S.; Shapiro, B.; Holmes, E.C. Insights into the evolutionary history of an emerging livestock pathogen: porcine circovirus 2. *J. Virol.* **2009**, *83*, 12813–12821.
69. Tischer, I.; Peters, D.; Rasch, R.; Pociuli, S. Replication of porcine circovirus: induction by glucosamine and cell cycle dependence. *Archives of Virology* **1987**, *96*, 39–57.
70. Dean, F.B.; Nelson, J.R.; Giesler, T.L.; Lasken, R.S. Rapid Amplification of Plasmid and Phage DNA Using Phi29 DNA Polymerase and Multiply-Primed Rolling Circle Amplification. *Genome Res.* **2001**, *11*, 1095–1099.
71. Binga, E.K.; Lasken, R.S.; Neufeld, J.D. Something from (almost) nothing: the impact of multiple displacement amplification on microbial ecology. *ISME J* **2008**, *2*, 233–241.
72. Ilyina, T.V.; Koonin, E.V. Conserved sequence motifs in the initiator proteins for rolling circle DNA replication encoded by diverse replicons from eubacteria, eucaryotes and archaeobacteria. *Nucleic Acids Res.* **1992**, *20*, 3279–3285.
73. Koonin, E.V. A common set of conserved motifs in a vast variety of putative nucleic acid-dependent ATPases including MCM proteins involved in the initiation of eukaryotic DNA replication. *Nucleic Acids Res.* **1993**, *21*, 2541–2547.
74. Gorbalenya, A.E.; Koonin, E.V.; Wolf, Y.I. A new superfamily of putative NTP-binding domains encoded by genomes of small DNA and RNA viruses. *FEBS Lett.* **1990**, *262*, 145–148.
75. Rybicki, E.P. A phylogenetic and evolutionary justification for three genera of Geminiviridae. *Arch. Virol.* **1994**, *139*, 49–77.
76. Biagini P; Bendinelli M; Hino S; Kakkola L; Mankertz A; Niel C; Okamoto H; Raidal S; Teo CG; Todd D Family Circoviridae. In *Virus taxonomy: classification and nomenclature of viruses: ninth report of the International Committee on Taxonomy of Viruses.*; Academic Press, London, 2012; pp. 343–349.
77. Tischer, I.; Gelderblom, H.; Vettermann, W.; Koch, M.A. A very small porcine virus with circular single-stranded DNA. *Nature* **1982**, *295*, 64–66.

78. Li, L.; Victoria, J.G.; Wang, C.; Jones, M.; Fellers, G.M.; Kunz, T.H.; Delwart, E. Bat guano virome: predominance of dietary viruses from insects and plants plus novel mammalian viruses. *J. Virol.* **2010**, *84*, 6955–6965.
79. Li, L.; Shan, T.; Soji, O.B.; Alam, M.M.; Kunz, T.H.; Zaidi, S.Z.; Delwart, E. Possible cross-species transmission of circoviruses and cycloviruses among farm animals. *J. Gen. Virol.* **2011**, *92*, 768–772.
80. Lima, F.E.S.; Cibulski, S.P.; Dall Bello, A.G.; Mayer, F.Q.; Witt, A.A.; Roehe, P.M.; d'Azevedo, P.A. A Novel Chiropteran Circovirus Genome Recovered from a Brazilian Insectivorous Bat Species. *Genome Announc* **2015**, *3*.
81. Wu, Z.; Yang, L.; Ren, X.; He, G.; Zhang, J.; Yang, J.; Qian, Z.; Dong, J.; Sun, L.; Zhu, Y.; et al. Deciphering the bat virome catalog to better understand the ecological diversity of bat viruses and the bat origin of emerging infectious diseases. *ISME J* **2016**, *10*, 609–620.
82. Li, L.; Kapoor, A.; Slikas, B.; Bamidele, O.S.; Wang, C.; Shaukat, S.; Masroor, M.A.; Wilson, M.L.; Ndjanga, J.-B.N.; Peeters, M.; et al. Multiple diverse circoviruses infect farm animals and are commonly found in human and chimpanzee feces. *J. Virol.* **2010**, *84*, 1674–1682.
83. Li, L.; McGraw, S.; Zhu, K.; Leutenegger, C.M.; Marks, S.L.; Kubiski, S.; Gaffney, P.; Dela Cruz, F.N.; Wang, C.; Delwart, E.; et al. Circovirus in tissues of dogs with vasculitis and hemorrhage. *Emerging Infect. Dis.* **2013**, *19*, 534–541.
84. Lian, H.; Liu, Y.; Li, N.; Wang, Y.; Zhang, S.; Hu, R. Novel circovirus from mink, China. *Emerging Infect. Dis.* **2014**, *20*, 1548–1550.
85. Phan, T.G.; Luchsinger, V.; Avendaño, L.F.; Deng, X.; Delwart, E. Cyclovirus in nasopharyngeal aspirates of Chilean children with respiratory infections. *Journal of General Virology* **2014**, *95*, 922–927.
86. Cui, L.; Wu, B.; Zhu, X.; Guo, X.; Ge, Y.; Zhao, K.; Qi, X.; Shi, Z.; Zhu, F.; Sun, L.; et al. Identification and genetic characterization of a novel circular single-stranded DNA virus in a human upper respiratory tract sample. *Arch Virol* **2017**, *162*, 3305–3312.
87. Tan, L.V.; van Doorn, H.R.; Nghia, H.D.T.; Chau, T.T.H.; Tu, L.T.P.; de Vries, M.; Canuti, M.; Deijs, M.; Jebbink, M.F.; Baker, S.; et al. Identification of a new cyclovirus in cerebrospinal fluid of patients with acute central nervous system infections. *MBio* **2013**, *4*, e00231-00213.
88. Phan, T.G.; Mori, D.; Deng, X.; Rajindrajith, S.; Ranawaka, U.; Fan Ng, T.F.; Bucardo-Rivera, F.; Orlandi, P.; Ahmed, K.; Delwart, E. Small circular single stranded DNA viral genomes in unexplained cases of human encephalitis, diarrhea, and in untreated sewage. *Virology* **2015**, *482*, 98–104.
89. Phan, T.G.; da Costa, A.C.; Del Valle Mendoza, J.; Bucardo-Rivera, F.; Nordgren, J.; O’Ryan, M.; Deng, X.; Delwart, E. The fecal virome of South and Central American children with diarrhea includes small circular DNA viral genomes of unknown origin. *Arch. Virol.* **2016**, *161*, 959–966.
90. Sasaki, M.; Orba, Y.; Ueno, K.; Ishii, A.; Moonga, L.; Hang’ombe, B.M.; Mweene, A.S.; Ito, K.; Sawa, H. Metagenomic analysis of the shrew enteric virome reveals novel viruses related to human stool-associated viruses. *J. Gen. Virol.* **2015**, *96*, 440–452.
91. Castrignano, S.B.; Nagasse-Sugahara, T.K.; Garrafa, P.; Monezi, T.A.; Barrella, K.M.; Mehnert, D.U. Identification of circo-like virus-Brazil genomic sequences in raw sewage from the metropolitan area of São Paulo: evidence of circulation two and three years after the first detection. *Mem. Inst. Oswaldo Cruz* **2017**, *112*, 175–181.
92. Halary, S.; Duraisamy, R.; Fancello, L.; Monteil-Bouchard, S.; Jardot, P.; Biagini, P.; Gouriet, F.; Raoult, D.; Desnues, C. Novel Single-Stranded DNA Circular Viruses in Pericardial Fluid of Patient with Recurrent Pericarditis. *Emerging Infect. Dis.* **2016**, *22*, 1839–1841.
93. Rosario, K.; Breitbart, M.; Harrach, B.; Segalés, J.; Delwart, E.; Biagini, P.; Varsani, A. Revisiting the taxonomy of the family Circoviridae: establishment of the genus Cyclovirus and removal of the genus Gyrovirus. *Arch. Virol.* **2017**, *162*, 1447–1463.
94. Nayar, G.P.; Hamel, A.; Lin, L. Detection and characterization of porcine circovirus associated with postweaning multisystemic wasting syndrome in pigs. *Can. Vet. J.* **1997**, *38*, 385–386.



95. Opriessnig, T.; Meng, X.-J.; Halbur, P.G. Porcine circovirus type 2 associated disease: update on current terminology, clinical manifestations, pathogenesis, diagnosis, and intervention strategies. *J. Vet. Diagn. Invest.* **2007**, *19*, 591–615.
96. Gillespie, J.; Opriessnig, T.; Meng, X.J.; Pelzer, K.; Buechner-Maxwell, V. Porcine circovirus type 2 and porcine circovirus-associated disease. *J. Vet. Intern. Med.* **2009**, *23*, 1151–1163.
97. Rose, N.; Opriessnig, T.; Grasland, B.; Jestin, A. Epidemiology and transmission of porcine circovirus type 2 (PCV2). *Virus Res.* **2012**, *164*, 78–89.
98. Smits, S.L.; Zijlstra, E.E.; van Hellemond, J.J.; Schapendonk, C.M.E.; Bodewes, R.; Schürch, A.C.; Haagmans, B.L.; Osterhaus, A.D.M.E. Novel Cyclovirus in Human Cerebrospinal Fluid, Malawi, 2010–2011. *Emerg Infect Dis* **2013**, *19*, 1511–1513.
99. Abbas, A.A.; Taylor, L.J.; Dothard, M.I.; Leiby, J.S.; Fitzgerald, A.S.; Khatib, L.A.; Collman, R.G.; Bushman, F.D. Redondoviridae, a Family of Small, Circular DNA Viruses of the Human Oro-Respiratory Tract Associated with Periodontitis and Critical Illness. *Cell Host Microbe* **2019**, *25*, 719–729.e4.
100. Kilbourne, E.D. Taxonomy and Comparative Virology of the Influenza Viruses. In *Influenza*; Kilbourne, E.D., Ed.; Springer US: Boston, MA, 1987; pp. 25–32 ISBN 978-1-4684-5239-6.
101. Hause, B.M.; Collin, E.A.; Liu, R.; Huang, B.; Sheng, Z.; Lu, W.; Wang, D.; Nelson, E.A.; Li, F. Characterization of a novel influenza virus in cattle and Swine: proposal for a new genus in the Orthomyxoviridae family. *MBio* **2014**, *5*, e00031-00014.
102. Grenfell, B.T.; Pybus, O.G.; Gog, J.R.; Wood, J.L.N.; Daly, J.M.; Mumford, J.A.; Holmes, E.C. Unifying the epidemiological and evolutionary dynamics of pathogens. *Science* **2004**, *303*, 327–332.
103. McCauley, J.W.; Hongo, S.; Kaverin, N.V.; Kochs, G.; Lamb, R.A.; Matrosovich, M.N.; Perez, D.R.; Palese, P.; Presti, R.M.; Rimstad, E.; et al. Orthomyxoviridae. *Virus Taxonomy, Ninth Report of the International Committee on Taxonomy of Viruses* **2012**, 749–761.
104. Rota, P.A.; Wallis, T.R.; Harmon, M.W.; Rota, J.S.; Kendal, A.P.; Nerome, K. Cocirculation of two distinct evolutionary lineages of influenza type B virus since 1983. *Virology* **1990**, *175*, 59–68.
105. Homma, M.; Ohyama, S.; Katagiri, S. Age distribution of the antibody to type C influenza virus. *Microbiol. Immunol.* **1982**, *26*, 639–642.
106. Manuguerra, J.C.; Hannoun, C.; Sáenz, M. del C.; Villar, E.; Cabezas, J.A. Sero-epidemiological survey of influenza C virus infection in Spain. *Eur. J. Epidemiol.* **1994**, *10*, 91–94.
107. Asha, K.; Kumar, B. Emerging Influenza D Virus Threat: What We Know so Far! *J Clin Med* **2019**, *8*.
108. Sistema de Vigilancia de la Gripe en España. Informe de Vigilancia de la Gripe en España. Temporada 2017-18.
109. Anderson, C.S.; McCall, P.R.; Stern, H.A.; Yang, H.; Topham, D.J. Antigenic cartography of H1N1 influenza viruses using sequence-based antigenic distance calculation. *BMC Bioinformatics* **2018**, *19*.
110. Webster, R.G.; Bean, W.J.; Gorman, O.T.; Chambers, T.M.; Kawaoka, Y. Evolution and ecology of influenza A viruses. *Microbiol. Rev.* **1992**, *56*, 152–179.
111. Seong, M.-W.; Cho, S.I.; Park, H.; Seo, S.H.; Lee, S.J.; Kim, E.-C.; Park, S.S. Genotyping Influenza Virus by Next-Generation Deep Sequencing in Clinical Specimens. *Ann Lab Med* **2016**, *36*, 255–258.
112. Xiao, Y.-L.; Ren, L.; Zhang, X.; Qi, L.; Kash, J.C.; Xiao, Y.; Wu, F.; Wang, J.; Taubenberger, J.K. Deep Sequencing of H7N9 Influenza A Viruses from 16 Infected Patients from 2013 to 2015 in Shanghai Reveals Genetic Diversity and Antigenic Drift. *mSphere* **2018**, *3*.
113. Barbezange, C.; Jones, L.; Blanc, H.; Isakov, O.; Celniker, G.; Enouf, V.; Shomron, N.; Vignuzzi, M.; van der Werf, S. Seasonal Genetic Drift of Human Influenza A Virus Quasispecies Revealed by Deep Sequencing. *Front Microbiol* **2018**, *9*, 2596.
114. Sheng, Z.; Liu, R.; Yu, J.; Ran, Z.; Newkirk, S.J.; An, W.; Li, F.; Wang, D. Identification and characterization of viral defective RNA genomes in influenza B virus. *J. Gen. Virol.* **2018**, *99*, 475–488.

115. Beerenwinkel, N.; Günthard, H.F.; Roth, V.; Metzner, K.J. Challenges and opportunities in estimating viral genetic diversity from next-generation sequencing data. *Front Microbiol* **2012**, *3*, 329.
116. Dinis, J.M.; Florek, N.W.; Fatola, O.O.; Moncla, L.H.; Mutschler, J.P.; Charlier, O.K.; Meece, J.K.; Belongia, E.A.; Friedrich, T.C. Deep Sequencing Reveals Potential Antigenic Variants at Low Frequencies in Influenza A Virus-Infected Humans. *J. Virol.* **2016**, *90*, 3355–3365.
117. Koboldt, D.C.; Steinberg, K.M.; Larson, D.E.; Wilson, R.K.; Mardis, E.R. The next-generation sequencing revolution and its impact on genomics. *Cell* **2013**, *155*, 27–38.
118. European Centre for Disease Prevention and Control ECDC strategic framework for the integration of molecular and genomic typing into European surveillance and multi-country outbreak investigations 2019–2021 2019.
119. IUCN SSC Bat Specialist Group Available online: <http://www.iucnbsg.org/> (accessed on May 14, 2018).
120. Calisher CH; Childs JE; Field HE; Holmes KV; Schountz T Bats: important reservoir hosts of emerging viruses. *Clinical Microbiology Review* **2006**, *19*(3):531-45.
121. Hayman, D.T.S. Bats as Viral Reservoirs. *Annual Review of Virology* **2016**, *3*, 77–99.
122. Mingo-Casas, P.; Sandonis, V.; Vázquez-Morón, S.; Berciano, J.M.; Juste, J.; Echevarría, J. *Rabies in Spain. A peculiarity in Eurasia*; 2017;
123. Vázquez-Morón, S.; Juste, J.; Ibáñez, C.; Berciano, J.M.; Echevarría, J.E. Phylogeny of European Bat Lyssavirus 1 in *Eptesicus isabellinus* Bats, Spain. *Emerg Infect Dis* **2011**, *17*, 520–523.
124. Vázquez-Morón, S.; Juste, J.; Ibáñez, C.; Ruiz-Villamor, E.; Avellón, A.; Vera, M.; Echevarría, J.E. Endemic Circulation of European Bat Lyssavirus Type 1 in Serotine Bats, Spain. *Emerg Infect Dis* **2008**, *14*, 1263–1266.
125. Juste, J.; Bilgin, R.; Muñoz, J.; Ibáñez, C. Mitochondrial DNA signatures at different spatial scales: from the effects of the Straits of Gibraltar to population structure in the meridional serotine bat (*Eptesicus isabellinus*). *Heredity (Edinb)* **2009**, *103*, 178–187.
126. Ceballos, N.A.; Morón, S.V.; Berciano, J.M.; Nicolás, O.; López, C.A.; Juste, J.; Nevado, C.R.; Setién, Á.A.; Echevarría, J.E. Novel Lyssavirus in Bat, Spain. *Emerg Infect Dis* **2013**, *19*, 793–795.
127. Negredo, A.; Palacios, G.; Vázquez-Morón, S.; González, F.; Dopazo, H.; Molero, F.; Juste, J.; Quetglas, J.; Savji, N.; de la Cruz Martínez, M.; et al. Discovery of an ebolavirus-like filovirus in europe. *PLoS Pathog.* **2011**, *7*, e1002304.
128. Falcón A; Vázquez-Morón S; Casas I; Aznar C; Ruiz G; Pozo F; Perez-Breña P; Juste J; Ibáñez C; Garin I; et al. Detection of alpha and betacoronaviruses in multiple Iberian bat species. *Archives of Virology* **2011**, *156*(10):1883–90.
129. Pozo F; Juste J; Vázquez-Morón S; Aznar-López C; Ibáñez C; Garin I; Aihartza J; Casas I; Tenorio A; Echevarría JE Identification of Novel Betaherpesviruses in Iberian Bats Reveals Parallel Evolution. *PLoS One* **2016**, *11*(12):e0169153.
130. Aznar-Lopez, C.; Vazquez-Moron, S.; Marston, D.A.; Juste, J.; Ibanez, C.; Berciano, J.M.; Salsamendi, E.; Aihartza, J.; Banyard, A.C.; McElhinney, L.; et al. Detection of rhabdovirus viral RNA in oropharyngeal swabs and ectoparasites of Spanish bats. *Journal of General Virology* **2013**, *94*, 69–75.
131. Drexler JF; Corman VM; Wegner T; Tateno AF; Zerbinati RM; Gloza-Rausch F; Seebens A; Müller MA; Drosten C Amplification of emerging viruses in a bat colony. *Emerging infectious diseases* **2011**, *17*(3):449-56.
132. Jánoska M; Vidovszky M; Molnár V; Liptovszky M; Harrach B; Benkö M Novel adenoviruses and herpesviruses detected in bats. *The Veterinary Journal* **2011**, *189*(1):118–21.
133. Tan, B.; Yang, X.-L.; Ge, X.-Y.; Peng, C.; Zhang, Y.-Z.; Zhang, L.-B.; Shi, Z.-L. Novel bat adenoviruses with an extremely large E3 gene. *Journal of General Virology* **2016**, *97*, 1625–1635.
134. Waruhiu C; Ommeh S; Obanda V; Agwanda B; Gakuya F; Ge XY; Yang XL; Wu LJ; Zohaib A; Hu B; et al. Molecular detection of viruses in Kenyan bats and discovery of novel astroviruses, caliciviruses and rotaviruses. *Virologica Sinica* **2017**, *32*(2):101-114.

135. Harrach B; Benkö M; Both G; Brown M; Davison AJ; Echavarría M; Hess M; Jones MS; Kajon A; Lehmkuhl AD; et al. Family Adenoviridae. In *Virus Taxonomy: Classification and Nomenclature of Viruses. Ninth report of the International Committee of Taxonomy of Viruses.*; Elsevier: San Diego, USA, 2011; pp. 125–141.
136. Maeda K; Hondo E; Terakawa J; Kiso Y; Nakaichi N; Endoh D; Sakai K; Morikawa S; Mizutani T Isolation of novel adenovirus from fruit bat (*Pteropus dasymallus yayeyamae*). *Emerging infectious diseases* **2008**, *14*(2):347–9.
137. Sonntag M; Mühldorfer K; Speck S; Wibbelt G; Kurth A New adenovirus in bats, Germany. *Emerging infectious diseases* **2009**, *15*(12):2052–5.
138. Kohl C; Vidovszky M; Mühldorfer K; Dabrowski PW; Radonic A; Nitsche A; Wibbelt G; Kurth A; Harrach B Genome analysis of bat adenovirus 2: indications of interspecies transmission. *Journal of Virology* **2012**, *86*:1888–92.
139. Li Y; Ge X; Zhang H; Zhou P; Zhu Y; Zhang Y; Yuan J; Wang LF; Shi Z Host range, prevalence, and genetic diversity of adenoviruses in bats. *Journal of Virology* **2010**, *84*(8):3889–97.
140. Lima FE; Cibulski SP; Elesbao F; Carnieli Junior P; Batista HB; Roehe PM; Franco AC First detection of adenovirus in the vampire bat (*Desmodus rotundus*) in Brazil. *Virus Genes* **2013**, *47*(2):378–81.
141. Vidovszky M; Kohl C; Boldogh S; Görföl T; Wibbelt G; Kurth A; Harrach B Random sampling of the Central European bat fauna reveals the existence of numerous hitherto unknown adenoviruses. *Acta Veterinaria Hungarica* **2015**, *63*(4), 508–25.
142. Tan B; Yang XL; Ge XY; Peng C; Liu HZ; Zhang YZ; Zhang LB; Shi ZL Novel bat adenoviruses with low G+C content shed new light on the evolution of adenoviruses. *Journal of General Virology* **2017**, *98*(4):739–748.
143. Baker, K.S.; Leggett, R.M.; Bexfield, N.H.; Alston, M.; Daly, G.; Todd, S.; Tachedjian, M.; Holmes, C.E.G.; Cramer, S.; Wang, L.-F.; et al. Metagenomic study of the viruses of African straw-coloured fruit bats: Detection of a chiropteran poxvirus and isolation of a novel adenovirus. *Virology* **2013**, *441*, 95–106.
144. Ogawa, H.; Kajihara, M.; Nao, N.; Shigeno, A.; Fujikura, D.; Hang'ombe, B.M.; Mweene, A.S.; Mutemwa, A.; Squarre, D.; Yamada, M.; et al. Characterization of a Novel Bat Adenovirus Isolated from Straw-Colored Fruit Bat (*Eidolon helvum*). *Viruses* **2017**, *9*, 371.
145. Vuren, P.J. van; Allam, M.; Wiley, M.R.; Ismail, A.; Storm, N.; Birkhead, M.; Markotter, W.; Palacios, G.; Paweska, J.T. A novel adenovirus isolated from the Egyptian fruit bat in South Africa is closely related to recent isolates from China. *Scientific Reports* **2018**, *8*, 9584.
146. Hackenbrack, N.; Rogers, M.B.; Ashley, R.E.; Keel, M.K.; Kubiski, S.V.; Bryan, J.A.; Ghedin, E.; Holmes, E.C.; Hafenstein, S.L.; Allison, A.B. Evolution and Cryo-electron Microscopy Capsid Structure of a North American Bat Adenovirus and Its Relationship to Other Mastadenoviruses. *J. Virol.* **2017**, *91*.
147. Hang, J.; Vento, T.J.; Norby, E.A.; Jarman, R.G.; Keiser, P.B.; Kuschner, R.A.; Binn, L.N. Adenovirus type 4 respiratory infections with a concurrent outbreak of coxsackievirus A21 among United States Army Basic Trainees, a retrospective viral etiology study using next-generation sequencing. *J. Med. Virol.* **2017**, *89*, 1387–1394.
148. Hage, E.; Dhingra, A.; Liebert, U.G.; Bergs, S.; Ganzenmueller, T.; Heim, A. Three novel, multiple recombinant types of species of human mastadenovirus D (HAdV-D 73, 74 & 75) isolated from diarrhoeal faeces of immunocompromised patients. *J. Gen. Virol.* **2017**.
149. Ogilvie, M. Molecular techniques should not now replace cell culture in diagnostic virology laboratories. *Rev. Med. Virol.* **2001**, *11*, 351–354.
150. AlGhounaim, M.; Xiao, Y.; Caya, C.; Papenburg, J. Diagnostic yield and clinical impact of routine cell culture for respiratory viruses among children with a negative multiplex RT-PCR result. *Journal of Clinical Virology* **2017**, *94*, 107–109.
151. Briese, T.; Kapoor, A.; Mishra, N.; Jain, K.; Kumar, A.; Jabado, O.J.; Lipkin, W.I. Virome Capture Sequencing Enables Sensitive Viral Diagnosis and Comprehensive Virome Analysis. *mBio* **2015**, *6*, e01491–15.

152. Paskey, A.C.; Frey, K.G.; Schroth, G.; Gross, S.; Hamilton, T.; Bishop-Lilly, K.A. Enrichment post-library preparation enhances the sensitivity of high-throughput sequencing-based detection and characterization of viruses from complex samples. *BMC Genomics* **2019**, *20*, 155.
153. O'Flaherty, B.M.; Li, Y.; Tao, Y.; Paden, C.R.; Queen, K.; Zhang, J.; Dinwiddie, D.L.; Gross, S.M.; Schroth, G.P.; Tong, S. Comprehensive viral enrichment enables sensitive respiratory virus genomic identification and analysis by next generation sequencing. *Genome Res.* **2018**, *28*, 869–877.
154. Zhang, Y.; Bergelson, J.M. Adenovirus Receptors. *J Virol* **2005**, *79*, 12125–12131.
155. Yu, Z.; Zeng, Z.; Zhang, J.; Pan, Y.; Chen, M.; Guo, Y.; Yu, N.; Chodosh, J.; Fu, N.; Che, X.; et al. Fatal Community-acquired Pneumonia in Children Caused by Re-emergent Human Adenovirus 7d Associated with Higher Severity of Illness and Fatality Rate. *Sci Rep* **2016**, *6*, 37216.
156. Dosin, G.; Aoun, F.; El Rassy, E.; Assi, T.; Lewalle, P.; Blanc, J.; van Velthoven, R.; Bron, D. Viral-induced Hemorrhagic Cystitis After Allogeneic Hematopoietic Stem Cell Transplant. *Clin Lymphoma Myeloma Leuk* **2017**, *17*, 438–442.
157. Lee, B.; Park, E.; Ha, J.; Ha, I.S.; Il Cheong, H.; Kang, H.G. Disseminated adenovirus infection in a 10-year-old renal allograft recipient. *Kidney Res Clin Pract* **2018**, *37*, 414–417.
158. Coiras, M. t.; Pérez-Breña, P.; García, M. I.; Casas, I. Simultaneous detection of influenza A, B, and C viruses, respiratory syncytial virus, and adenoviruses in clinical samples by multiplex reverse transcription nested-PCR assay. *J. Med. Virol.* **2003**, *69*, 132–144.
159. Coiras, M. t.; Aguilar, J. c.; García, M. I.; Casas, I.; Pérez-Breña, P. Simultaneous detection of fourteen respiratory viruses in clinical specimens by two multiplex reverse transcription nested-PCR assays. *J. Med. Virol.* **2004**, *72*, 484–495.
160. Corman, V.M.; Eckerle, I.; Bleicker, T.; Zaki, A.; Landt, O.; Eschbach-Bludau, M.; van Boheemen, S.; Gopal, R.; Ballhause, M.; Bestebroer, T.M.; et al. Detection of a novel human coronavirus by real-time reverse-transcription polymerase chain reaction. *Euro Surveill.* **2012**, *17*.
161. Casas, I.; Powell, L.; Klapper, P.E.; Cleator, G.M. New method for the extraction of viral RNA and DNA from cerebrospinal fluid for use in the polymerase chain reaction assay. *Journal of Virological Methods* **1995**, *53*, 25–36.
162. Nextera XT DNA Library Prep Kit Reference Guide (15031942). *Illumina Corporation*. **2019**
163. Jiang, H.; Lei, R.; Ding, S.-W.; Zhu, S. Skewer: a fast and accurate adapter trimmer for next-generation sequencing paired-end reads. *BMC Bioinformatics* **2014**, *15*, 182.
164. Peng, Y.; Leung, H.C.M.; Yiu, S.M.; Chin, F.Y.L. IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics* **2012**, *28*, 1420–1428.
165. Altschul, S.F.; Madden, T.L.; Schäffer, A.A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D.J. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **1997**, *25*, 3389–3402.
166. Zhao, Y.; Tang, H.; Ye, Y. RAPSearch2: a fast and memory-efficient protein similarity search tool for next-generation sequencing data. *Bioinformatics* **2012**, *28*, 125–126.
167. Coil, D.; Jospin, G.; Darling, A.E. A5-miseq: an updated pipeline to assemble microbial genomes from Illumina MiSeq data. *Bioinformatics* **2015**, *31*, 587–589.
168. Kajitani, R.; Toshimoto, K.; Noguchi, H.; Toyoda, A.; Ogura, Y.; Okuno, M.; Yabana, M.; Harada, M.; Nagayasu, E.; Maruyama, H.; et al. Efficient de novo assembly of highly heterozygous genomes from whole-genome shotgun short reads. *Genome Res.* **2014**, *24*, 1384–1395.
169. Ruby, J.G.; Bellare, P.; Derisi, J.L. PRICE: software for the targeted assembly of components of (Meta) genomic sequence data. *G3 (Bethesda)* **2013**, *3*, 865–880.
170. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **2011**, *17*, 10–12.
171. Schmieder, R.; Edwards, R. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* **2011**, *27*, 863–864.
172. Langmead, B.; Trapnell, C.; Pop, M.; Salzberg, S.L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology* **2009**, *10*, R25.

173. Boisvert, S.; Raymond, F.; Godzaridis, É.; Laviolette, F.; Corbeil, J. Ray Meta: scalable de novo metagenome assembly and profiling. *Genome Biol* **2012**, *13*, R122.
174. Buchfink, B.; Xie, C.; Huson, D.H. Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* **2015**, *12*, 59–60.
175. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120.
176. Bankevich, A.; Nurk, S.; Antipov, D.; Gurevich, A.A.; Dvorkin, M.; Kulikov, A.S.; Lesin, V.M.; Nikolenko, S.I.; Pham, S.; Prjibelski, A.D.; et al. SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *J Comput Biol* **2012**, *19*, 455–477.
177. Shepard, S.S.; Davis, C.T.; Bahl, J.; Rivaller, P.; York, I.A.; Donis, R.O. LABEL: Fast and Accurate Lineage Assignment with Assessment of H5N1 and H9N2 Influenza A Hemagglutinins. *PLOS ONE* **2014**, *9*, e86921.
178. Posada, D.; Crandall, K.A. MODELTEST: testing the model of DNA substitution. *Bioinformatics* **1998**, *14*, 817–818.
179. Reyes, A.; Haynes, M.; Hanson, N.; Angly, F.E.; Heath, A.C.; Rohwer, F.; Gordon, J.I. Viruses in the faecal microbiota of monozygotic twins and their mothers. *Nature* **2010**, *466*, 334–338.
180. Daly, G.M.; Bexfield, N.; Heaney, J.; Stubbs, S.; Mayer, A.P.; Palser, A.; Kellam, P.; Drou, N.; Caccamo, M.; Tiley, L.; et al. A Viral Discovery Methodology for Clinical Biopsy Samples Utilising Massively Parallel Next Generation Sequencing. *PLOS ONE* **2011**, *6*, e28879.
181. Thurber, R.V.; Haynes, M.; Breitbart, M.; Wegley, L.; Rohwer, F. Laboratory procedures to generate viral metagenomes. *Nat Protoc* **2009**, *4*, 470–483.
182. Minot, S.; Sinha, R.; Chen, J.; Li, H.; Keilbaugh, S.A.; Wu, G.D.; Lewis, J.D.; Bushman, F.D. The human gut virome: inter-individual variation and dynamic response to diet. *Genome Res.* **2011**, *21*, 1616–1625.
183. Pride, D.T.; Salzman, J.; Haynes, M.; Rohwer, F.; Davis-Long, C.; White, R.A.; Loomer, P.; Armitage, G.C.; Relman, D.A. Evidence of a robust resident bacteriophage population revealed through analysis of the human salivary virome. *ISME J* **2012**, *6*, 915–926.
184. Hannigan, G.D.; Meisel, J.S.; Tyldsley, A.S.; Zheng, Q.; Hodgkinson, B.P.; SanMiguel, A.J.; Minot, S.; Bushman, F.D.; Grice, E.A. The human skin double-stranded DNA virome: topographical and temporal diversity, genetic enrichment, and dynamic associations with the host microbiome. *MBio* **2015**, *6*, e01578-01515.
185. Lim, E.S.; Zhou, Y.; Zhao, G.; Bauer, I.K.; Droit, L.; Ndao, I.M.; Warner, B.B.; Tarr, P.I.; Wang, D.; Holtz, L.R. Early life dynamics of the human gut virome and bacterial microbiome in infants. *Nat. Med.* **2015**, *21*, 1228–1234.
186. Aggarwala, V.; Liang, G.; Bushman, F.D. Viral communities of the human gut: metagenomic analysis of composition and dynamics. *Mob DNA* **2017**, *8*, 12.
187. Victoria, J.G.; Kapoor, A.; Li, L.; Blinkova, O.; Slikas, B.; Wang, C.; Naeem, A.; Zaidi, S.; Delwart, E. Metagenomic analyses of viruses in stool samples from children with acute flaccid paralysis. *J. Virol.* **2009**, *83*, 4642–4651.
188. Karlsson, O.E.; Belák, S.; Granberg, F. The effect of preprocessing by sequence-independent, single-primer amplification (SISPA) on metagenomic detection of viruses. *Biosecur Bioterror* **2013**, *11 Suppl 1*, S227–234.
189. Kleiner, M.; Hooper, L.V.; Duerkop, B.A. Evaluation of methods to purify virus-like particles for metagenomic sequencing of intestinal viromes. *BMC Genomics* **2015**, *16*, 7.
190. Castro-Mejía, J.L.; Muhammed, M.K.; Kot, W.; Neve, H.; Franz, C.M.A.P.; Hansen, L.H.; Vogensen, F.K.; Nielsen, D.S. Optimizing protocols for extraction of bacteriophages prior to metagenomic analyses of phage communities in the human gut. *Microbiome* **2015**, *3*, 64.
191. Lewandowska, D.W.; Zagordi, O.; Geissberger, F.-D.; Kufner, V.; Schmutz, S.; Böni, J.; Metzner, K.J.; Trkola, A.; Huber, M. Optimization and validation of sample preparation for metagenomic sequencing of viruses in clinical samples. *Microbiome* **2017**, *5*, 94.

192. Parras-Moltó, M.; Rodríguez-Galet, A.; Suárez-Rodríguez, P.; López-Bueno, A. Evaluation of bias induced by viral enrichment and random amplification protocols in metagenomic surveys of saliva DNA viruses. *Microbiome* **2018**, *6*.
193. López-Bueno, A.; Tamames, J.; Velázquez, D.; Moya, A.; Quesada, A.; Alcamí, A. High diversity of the viral community from an Antarctic lake. *Science* **2009**, *326*, 858–861.
194. Li, L.; Deng, X.; Mee, E.T.; Collot-Teixeira, S.; Anderson, R.; Schepelmann, S.; Minor, P.D.; Delwart, E. Comparing viral metagenomics methods using a highly multiplexed human viral pathogens reagent. *J. Virol. Methods* **2015**, *213*, 139–146.
195. Conceição-Neto, N.; Zeller, M.; Lefrère, H.; De Bruyn, P.; Beller, L.; Deboutte, W.; Yinda, C.K.; Lavigne, R.; Maes, P.; Van Ranst, M.; et al. Modular approach to customise sample preparation procedures for viral metagenomics: a reproducible protocol for virome analysis. *Sci Rep* **2015**, *5*, 16532.
196. Wendl, M.C. Occupancy modeling of coverage distribution for whole genome shotgun DNA sequencing. *Bull. Math. Biol.* **2006**, *68*, 179–196.
197. Rosseel, T.; Van Borm, S.; Vandenbussche, F.; Hoffmann, B.; van den Berg, T.; Beer, M.; Höper, D. The origin of biased sequence depth in sequence-independent nucleic acid amplification and optimization for efficient massive parallel sequencing. *PLoS ONE* **2013**, *8*, e76144.
198. Atmar, R.L.; Piedra, P.A.; Patel, S.M.; Greenberg, S.B.; Couch, R.B.; Glezen, W.P. Picornavirus, the Most Common Respiratory Virus Causing Infection among Patients of All Ages Hospitalized with Acute Respiratory Illness. *Journal of Clinical Microbiology* **2012**, *50*, 506–508.
199. Simmonds, P.; McIntyre, C.; Savolainen-Kopra, C.; Tapparel, C.; Mackay, I.M.; Hovi, T. Proposals for the classification of human rhinovirus species C into genotypically assigned types. *J. Gen. Virol.* **2010**, *91*, 2409–2419.
200. Zell, R.; Delwart, E.; Gorbalenya, A.E.; Hovi, T.; King, A.M.Q.; Knowles, N.J.; Lindberg, A.M.; Pallansch, M.A.; Palmenberg, A.C.; Reuter, G.; et al. ICTV Virus Taxonomy Profile: Picornaviridae. *J. Gen. Virol.* **2017**, *98*, 2421–2422.
201. Gaynor, A.M.; Nissen, M.D.; Whitley, D.M.; Mackay, I.M.; Lambert, S.B.; Wu, G.; Brennan, D.C.; Storch, G.A.; Sloots, T.P.; Wang, D. Identification of a novel polyomavirus from patients with acute respiratory tract infections. *PLoS Pathog.* **2007**, *3*, e64.
202. Norja, P.; Ubillos, I.; Templeton, K.; Simmonds, P. No evidence for an association between infections with WU and KI polyomaviruses and respiratory disease. *J. Clin. Virol.* **2007**, *40*, 307–311.
203. Rao, S.; Lucero, M.G.; Nohynek, H.; Tallo, V.; Lupisan, S.P.; Garcea, R.L.; Simões, E.A.F.; ARIVAC Consortium WU and KI polyomavirus infections in Filipino children with lower respiratory tract disease. *J. Clin. Virol.* **2016**, *82*, 112–118.
204. Wang, Y.; Zhu, N.; Li, Y.; Lu, R.; Wang, H.; Liu, G.; Zou, X.; Xie, Z.; Tan, W. Metagenomic analysis of viral genetic diversity in respiratory samples from children with severe acute respiratory infection in China. *Clin. Microbiol. Infect.* **2016**, *22*, 458.e1–9.
205. Gillison, M.L.; Alemany, L.; Snijders, P.J.F.; Chaturvedi, A.; Steinberg, B.M.; Schwartz, S.; Castellsagué, X. Human papillomavirus and diseases of the upper airway: head and neck cancer and respiratory papillomatosis. *Vaccine* **2012**, *30 Suppl 5*, F34–54.
206. Nguyen, K.D.; Lee, E.E.; Yue, Y.; Stork, J.; Pock, L.; North, J.P.; Vandergriff, T.; Cockerell, C.; Hosler, G.A.; Pastrana, D.V.; et al. Human Polyomavirus 6 and 7 Are Associated with a Pruritic and Dyskeratotic Dermatitis. *J Am Acad Dermatol* **2017**, *76*, 932–940.e3.
207. Brealey, J.C.; Sly, P.D.; Young, P.R.; Chappell, K.J. Viral bacterial co-infection of the respiratory tract during early childhood. *FEMS Microbiol. Lett.* **2015**, *362*.
208. Noviello, S.; Huang, D.B. The Basics and the Advancements in Diagnosis of Bacterial Lower Respiratory Tract Infections. *Diagnostics (Basel)* **2019**, *9*.
209. Advani, S.; Sengupta, A.; Forman, M.; Valsamakis, A.; Milstone, A.M. Detecting respiratory viruses in asymptomatic children. *Pediatr. Infect. Dis. J.* **2012**, *31*, 1221–1226.
210. Falsey, A.R. Human metapneumovirus infection in adults. *Pediatr. Infect. Dis. J.* **2008**, *27*, S80–83.

211. Falsey, A.R.; Erdman, D.; Anderson, L.J.; Walsh, E.E. Human metapneumovirus infections in young and elderly adults. *J. Infect. Dis.* **2003**, *187*, 785–790.
212. Zhou, Y.; Fernandez, S.; Yoon, I.-K.; Simasathien, S.; Watanaveeradej, V.; Yang, Y.; Marte-Salcedo, O.A.; Shuck-Lee, D.J.; Thomas, S.J.; Hang, J.; et al. Metagenomics Study of Viral Pathogens in Undiagnosed Respiratory Specimens and Identification of Human Enteroviruses at a Thailand Hospital. *Am. J. Trop. Med. Hyg.* **2016**, *95*, 663–669.
213. Taboada, B.; Espinoza, M.A.; Isa, P.; Aponte, F.E.; Arias-Ortiz, M.A.; Monge-Martínez, J.; Rodríguez-Vázquez, R.; Díaz-Hernández, F.; Zárate-Vidal, F.; Wong-Chew, R.M.; et al. Is there still room for novel viral pathogens in pediatric respiratory tract infections? *PLoS ONE* **2014**, *9*, e113570.
214. Xu, L.; Zhu, Y.; Ren, L.; Xu, B.; Liu, C.; Xie, Z.; Shen, K. Characterization of the nasopharyngeal viral microbiome from children with community-acquired pneumonia but negative for Luminex xTAG respiratory viral panel assay detection. *J. Med. Virol.* **2017**, *89*, 2098–2107.
215. Vandini, S.; Biagi, C.; Fischer, M.; Lanari, M. Impact of Rhinovirus Infections in Children. *Viruses* **2019**, *11*.
216. Jacobs, S.E.; Lamson, D.M.; St George, K.; Walsh, T.J. Human rhinoviruses. *Clin. Microbiol. Rev.* **2013**, *26*, 135–162.
217. Howard, L.M.; Johnson, M.; Gil, A.I.; Griffin, M.R.; Edwards, K.M.; Lanata, C.F.; Williams, J.V.; Grijalva, C.G. Molecular Epidemiology of Rhinovirus Detections in Young Children. *Open Forum Infect Dis* **2016**, *3*, ofw001.
218. Rahamat-Langendoen, J.C.; Riezebos-Brilman, A.; Hak, E.; Schölvink, E.H.; Niesters, H.G.M. The significance of rhinovirus detection in hospitalized children: clinical, epidemiological and virological features. *Clin. Microbiol. Infect.* **2013**, *19*, E435–442.
219. Gwaltney, J.M.; Hendley, J.O.; Simon, G.; Jordan, W.S. Rhinovirus infections in an industrial population. I. The occurrence of illness. *N. Engl. J. Med.* **1966**, *275*, 1261–1268.
220. Lau, S.K.P.; Yip, C.C.Y.; Woo, P.C.Y.; Yuen, K.-Y. Human rhinovirus C: a newly discovered human rhinovirus species. *Emerg Health Threats J* **2010**, *3*, e2.
221. Jartti, T.; Lehtinen, P.; Vuorinen, T.; Koskenvuo, M.; Ruuskanen, O. Persistence of rhinovirus and enterovirus RNA after acute respiratory illness in children. *J. Med. Virol.* **2004**, *72*, 695–699.
222. Laine, P.; Blomqvist, S.; Savolainen, C.; Andries, K.; Hovi, T. Alignment of capsid protein VP1 sequences of all human rhinovirus prototype strains: conserved motifs and functional domains. *Journal of General Virology*, **2006**, *87*, 129–138.
223. Drake, J.W.; Holland, J.J. Mutation rates among RNA viruses. *PNAS* **1999**, *96*, 13910–13913.
224. Waman, V.P.; Kolekar, P.S.; Kale, M.M.; Kulkarni-Kale, U. Population Structure and Evolution of Rhinoviruses. *PLOS ONE* **2014**, *9*, e88981.
225. Palmenberg, A.C.; Spiro, D.; Kuzmickas, R.; Wang, S.; Djikeng, A.; Rathe, J.A.; Fraser-Liggett, C.M.; Liggett, S.B. Sequencing and Analyses of All Known Human Rhinovirus Genomes Reveal Structure and Evolution. *Science* **2009**, *324*, 55–59.
226. Canducci, F.; Debiaggi, M.; Sampaolo, M.; Marinozzi, M.C.; Berrè, S.; Terulla, C.; Gargantini, G.; Cambieri, P.; Romero, E.; Clementi, M. Two-year prospective study of single infections and co-infections by respiratory syncytial virus and viruses identified recently in infants with acute respiratory disease. *J. Med. Virol.* **2008**, *80*, 716–723.
227. Tanner, H.; Boxall, E.; Osman, H. Respiratory viral infections during the 2009–2010 winter season in Central England, UK: incidence and patterns of multiple virus co-infections. *Eur. J. Clin. Microbiol. Infect. Dis.* **2012**, *31*, 3001–3006.
228. Calvo, C.; García-garcía, M.L.; Pozo, F.; Paula, G.; Molinero, M.; Calderón, A.; González-esquivillas, M.; Casas, I. Respiratory Syncytial Virus Coinfections With Rhinovirus and Human Bocavirus in Hospitalized Children. *Medicine* **2015**, *94*.
229. Maggi, F.; Focosi, D.; Albani, M.; Lanini, L.; Vatteroni, M.L.; Petrini, M.; Ceccherini-Nelli, L.; Pistello, M.; Bendinelli, M. Role of hematopoietic cells in the maintenance of chronic human torquetenovirus plasma viremia. *J. Virol.* **2010**, *84*, 6891–6893.

230. Focosi, D.; Antonelli, G.; Pistello, M.; Maggi, F. Torquetenovirus: the human virome from bench to bedside. *Clin. Microbiol. Infect.* **2016**, *22*, 589–593.
231. Freer, G.; Maggi, F.; Pifferi, M.; Di Cicco, M.E.; Peroni, D.G.; Pistello, M. The Virome and Its Major Component, Anellovirus, a Convolved System Molding Human Immune Defenses and Possibly Affecting the Development of Asthma and Respiratory Diseases in Childhood. *Front Microbiol* **2018**, *9*, 686.
232. Maggi, F.; Bendinelli, M. Immunobiology of the Torque teno viruses and other anelloviruses. *Curr. Top. Microbiol. Immunol.* **2009**, *331*, 65–90.
233. Zhao, L.; Rosario, K.; Breitbart, M.; Duffy, S. Eukaryotic Circular Rep-Encoding Single-Stranded DNA (CRESS DNA) Viruses: Ubiquitous Viruses With Small Genomes and a Diverse Host Range. *Adv. Virus Res.* **2019**, *103*, 71–133.
234. Chandra, R. Picobirnavirus, a novel group of undescribed viruses of mammals and birds: a minireview. *Acta Virol.* **1997**, *41*, 59–62.
235. Grinde, B. Herpesviruses: latency and reactivation - viral strategies and host response. *J Oral Microbiol* **2013**, *5*.
236. Naccache, S.N.; Greninger, A.L.; Lee, D.; Coffey, L.L.; Phan, T.; Rein-Weston, A.; Aronsohn, A.; Hackett, J.; Delwart, E.L.; Chiu, C.Y. The perils of pathogen discovery: origin of a novel parvovirus-like hybrid genome traced to nucleic acid extraction spin columns. *J. Virol.* **2013**, *87*, 11966–11977.
237. Daudt, C.; da Silva, F.R.C.; Streck, A.F.; Weber, M.N.; Mayer, F.Q.; Cibulski, S.P.; Canal, C.W. How many papillomavirus species can go undetected in papilloma lesions? *Sci Rep* **2016**, *6*, 36480.
238. Rosario, K.; Duffy, S.; Breitbart, M. A field guide to eukaryotic circular single-stranded DNA viruses: insights gained from metagenomics. *Arch Virol* **2012**, *157*, 1851–1871.
239. Buller, R.M.; Rose, J.A. Characterization of adenovirus-associated virus-induced polypeptides in KB cells. *J. Virol.* **1978**, *25*, 331–338.
240. Wylie, K.M.; Mihindukulasuriya, K.A.; Sodergren, E.; Weinstock, G.M.; Storch, G.A. Sequence analysis of the human virome in febrile and afebrile children. *PLoS ONE* **2012**, *7*, e27735.
241. Rodríguez, J.M.; Murphy, K.; Stanton, C.; Ross, R.P.; Kober, O.I.; Juge, N.; Avershina, E.; Rudi, K.; Narbad, A.; Jenmalm, M.C.; et al. The composition of the gut microbiota throughout life, with an emphasis on early life. *Microb Ecol Health Dis* **2015**, *26*.
242. Rothberg, J.M.; Hinz, W.; Rearick, T.M.; Schultz, J.; Mileski, W.; Davey, M.; Leamon, J.H.; Johnson, K.; Milgrew, M.J.; Edwards, M.; et al. An integrated semiconductor device enabling non-optical genome sequencing. *Nature* **2011**, *475*, 348–352.
243. Eid, J.; Fehr, A.; Gray, J.; Luong, K.; Lyle, J.; Otto, G.; Peluso, P.; Rank, D.; Baybayan, P.; Bettman, B.; et al. Real-time DNA sequencing from single polymerase molecules. *Science* **2009**, *323*, 133–138.
244. Bentley, D.R.; Balasubramanian, S.; Swerdlow, H.P.; Smith, G.P.; Milton, J.; Brown, C.G.; Hall, K.P.; Evers, D.J.; Barnes, C.L.; Bignell, H.R.; et al. Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* **2008**, *456*, 53–59.
245. Simpson, J.T.; Workman, R.E.; Zuzarte, P.C.; David, M.; Dursi, L.J.; Timp, W. Detecting DNA cytosine methylation using nanopore sequencing. *Nature Methods* **2017**, *14*, 407–410.
246. Tyler, A.D.; Mataseje, L.; Urfano, C.J.; Schmidt, L.; Antonation, K.S.; Mulvey, M.R.; Corbett, C.R. Evaluation of Oxford Nanopore's MinION Sequencing Device for Microbial Whole Genome Sequencing Applications. *Scientific Reports* **2018**, *8*, 10931.
247. Wang, J.; Moore, N.E.; Deng, Y.-M.; Eccles, D.A.; Hall, R.J. MinION nanopore sequencing of an influenza genome. *Front Microbiol* **2015**, *6*, 766.
248. Seitz, V.; Schaper, S.; Dröge, A.; Lenze, D.; Hummel, M.; Hennig, S. A new method to prevent carry-over contaminations in two-step PCR NGS library preparations. *Nucleic Acids Res* **2015**, *43*, e135–e135.
249. Iglesias-Caballero, M.; Juste, J.; Vázquez-Morón, S.; Falcon, A.; Aznar-Lopez, C.; Ibáñez, C.; Pozo, F.; Ruiz, G.; Berciano, J.M.; Garin, I.; et al. New Adenovirus Groups in Western Palaearctic Bats. *Viruses* **2018**, *10*, 443.



250. García-Mudarra, J.L.; Ibáñez, C.; Juste, J. The Straits of Gibraltar: barrier or bridge to Ibero-Moroccan bat diversity? *Biological Journal of the Linnean Society* **2009**, *96*, 434–450.
251. Choi, K.H. Viral Polymerases. *Adv Exp Med Biol* **2012**, *726*, 267–304.
252. Roberts DM; Nanda A; Havenga MJ; Abbink P; Lynch DM; Ewald BA; Liu J; Thorner AR; Swanson PE; Gorgone DA; et al. Hexon-chimaeric adenovirus serotype 5 vectors circumvent pre-existing anti-vector immunity. *Nature* **2006**, *441*(7090):239–43.
253. Rux, J.J.; Kuser, P.R.; Burnett, R.M. Structural and phylogenetic analysis of adenovirus hexons by use of high-resolution x-ray crystallographic, molecular modeling, and sequence-based methods. *J. Virol.* **2003**, *77*, 9553–9566.
254. Calvo, C.; García-García, M.L.; Sanchez-Dehesa, R.; Román, C.; Tabares, A.; Pozo, F.; Casas, I. Eight Year Prospective Study of Adenoviruses Infections in Hospitalized Children. Comparison with Other Respiratory Viruses. *PLoS ONE* **2015**, *10*, e0132162.
255. Kajon, A.E.; Dickson, L.M.; Murtagh, P.; Viale, D.; Carballal, G.; Echavarria, M. Molecular Characterization of an Adenovirus 3-16 Intertypic Recombinant Isolated in Argentina from an Infant Hospitalized with Acute Respiratory Infection. *J Clin Microbiol* **2010**, *48*, 1494–1496.
256. Arroyo-Cabrales J; Álvarez-Castañeda ST *Corynorhinus rafinesquii*. The IUCN Red List of Threatened Species 2017 2017.
257. Stadelmann, B.; Lin, L.-K.; Kunz, T.H.; Ruedi, M. Molecular phylogeny of New World Myotis (Chiroptera, Vespertilionidae) inferred from mitochondrial and nuclear DNA genes. *Molecular Phylogenetics and Evolution* **2007**, *43*, 32–48.
258. Lun, J.H.; Crosbie, N.D.; White, P.A. Genetic diversity and quantification of human mastadenoviruses in wastewater from Sydney and Melbourne, Australia. *Science of The Total Environment* **2019**, *675*, 305–312.
259. Latham, G. DNase I demystified. *Thermo Fisher*. **2011**.

# Congresos y artículos



## ADENOVIRUS GROUPS IN WESTERN PALEARCTIC BATS IN SPAIN

Iglesias-Caballero M.<sup>1</sup>, Vázquez-Morón S.<sup>1</sup>, Juste J.<sup>2</sup>, Falcon A.<sup>1</sup>, Aznar C.<sup>1</sup>, Ibáñez C.<sup>2</sup>, Pozo F.<sup>1</sup>, Ruiz G.<sup>1</sup>, Berciano JM.<sup>1</sup>, Echevarría JE.<sup>1</sup>, Casas I.<sup>1</sup>.

(1)National Center for Microbiology, ISCIII, (2)Estación Biológica de Doñana. Consejo Superior de Investigaciones Científicas (CSIC)

**Objective:** General screening for adenoviruses in different spanish bat species (BtAdV).

This study is a long-term screening for AdV in a wide group of Western Palearctic bats with the purpose of identifying the taxonomic identities of the spanish BtAdV.

### Introduction:

- The first published BtAdV1 was isolated from a megachiropteran (*Pteropus dasymallus yaeyamae*) in 2008.
- In 2010, the first fully sequenced genome from a microchiropteran bat (*Myotis ricketti*) from China was published: *bat mastadenovirus A* specie.
- A BtAdV was isolated from a microchiropteran host (*Pipistrellus pipistrellus*) in Germany in 2009.
- In 2012, full genome established a second specie *bat mastadenovirus B*
- Several further BtAdVs have been detected by PCR in China, Hungary, Germany and by metagenomic study of guano from American bats.



Table 2

### Materials and methods:

- From 2004 to 2017 a total of 1395 bats belong to 27 different bat species (Table 1) were captured in several regions of Spain (Table 2).
- Bat were trapped in their shelters entrance.
- Samples were: oropharyngeal samples "OPS" (1392) and stool samples "SS" (325) (Table 1)
- Two different generic nested PCR protocols were: **DNA-dependent DNA polymerase** and **hexon** genes.

BAT SPECIES	OPS	SS
<i>Barbastella barbastellus</i> (Bba)	38	4
<i>Eptesicus isabellinus</i> (Eie)	0	8
<i>Eptesicus serotinus</i> (Ese)	0	14
<i>Hypsugo savii</i> (Hsa)	31	26 (3)
<i>Miniotropus schreibersii</i> (Msc)	152 (1)	2
<i>Myotis alcathoe</i> (Mal)	0	1
<i>Myotis bechsteinii</i> (Mbe)	18 (1)	2
<i>Myotis blythii</i> (Mbl)	29	0
<i>Myotis capaccini</i> (Mca)	15	0
<i>Myotis daubentonii</i> (Mda)	63	41
<i>Myotis emarginatus</i> (Mem)	56 (2)	0
<i>Myotis escaleri</i> (Mes)	13	0
<i>Myotis myotis</i> (Mmy)	79 (1)	1
<i>Myotis mystacinus</i> (Mmt)	2	8
<i>Myotis nattereri</i> (Mna)	36	3
<i>Nyctalus noctula</i> (Nno)	122 (3)	0
<i>Nyctalus lasiopterus</i> (Nly)	139 (9)	40 (7)
<i>Nyctalus leisleri</i> (Nle)	19 (1)	26 (3)
<i>Pipistrellus kuhlii</i> (Pku)	350 (11)	4 (3)
<i>Pipistrellus pipistrellus</i> (Ppi)	29 (1)	4 (1)
<i>Pipistrellus pygmaeus</i> (Ppy)	36 (7)	120 (12)
<i>Plecotus auritus</i> (Pau)	11	8
<i>Plecotus austriacus</i> (Pla)	10	6
<i>Rhinolophus euryale</i> (Rhe)	49 (6)	0
<i>Rhinolophus ferrumequinum</i> (Rfe)	90 (7)	3 (1)
<i>Rhinolophus hipposideros</i> (Rhi)	4	4
<i>Rhinolophus mehelyi</i> (Rme)	1	0
<b>TOTAL</b>	<b>1392 (38)</b>	<b>325 (28)</b>

Table 1

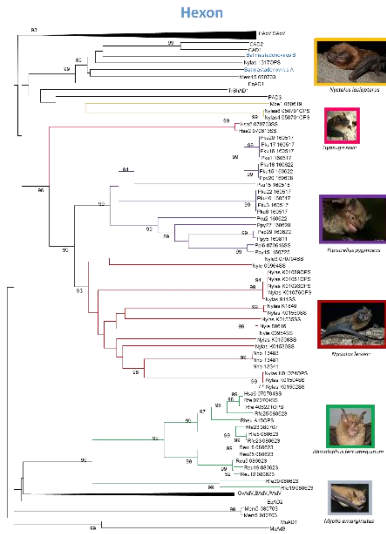


Fig. 1.a

### Results and conclusions:

- BtAdV was found in 66 bat samples: 38 OPS and 28 SS presence in both respiratory and intestinal tracts.
- AdVs in the Western Palearctic region show a wide range of bat hosts.
- Phylogenetic analyses on the partial sequences of the **hexon** and **DNA-dependent DNA polymerase** genes placed all Spanish BtAdVs within the genus *Mastadenovirus* (Figure 1).
- The clusters according the hexon (Figure 1a) and DNA-dependent DNA polymerase (Figure 2) genes were:

	Hexon	Polymerase
Mem 15-080703 clustered <i>Mastadenovirus A</i>	No	Yes
Ppy 14,15,21-160725 clustered with <i>Mastadenovirus B</i>	No	Yes
<b>Nyctalus Group 1</b>	Yes	Yes
Nyles 4,8-050701OPS as an independent group associated with Mbe1-080619: <b>Nyctalus group 2</b>	Yes	Yes
<b>Pipistrellus Group</b>	Yes	Yes
<b>Rhinolophus Group</b> associated with Hsa6-070704F	Yes	-
<b>Hypsugo Group</b>	Yes	-
<b>Myotis Group</b>	Yes	-

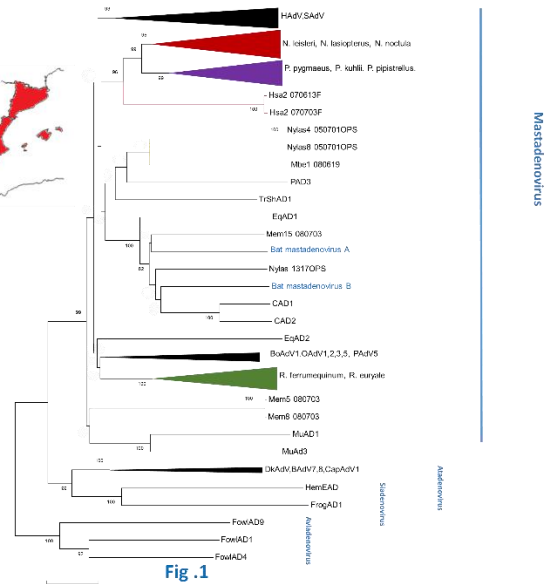


Fig. 1

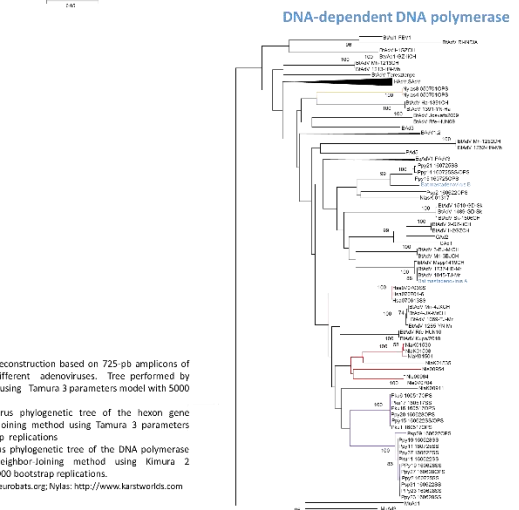


Fig. 2

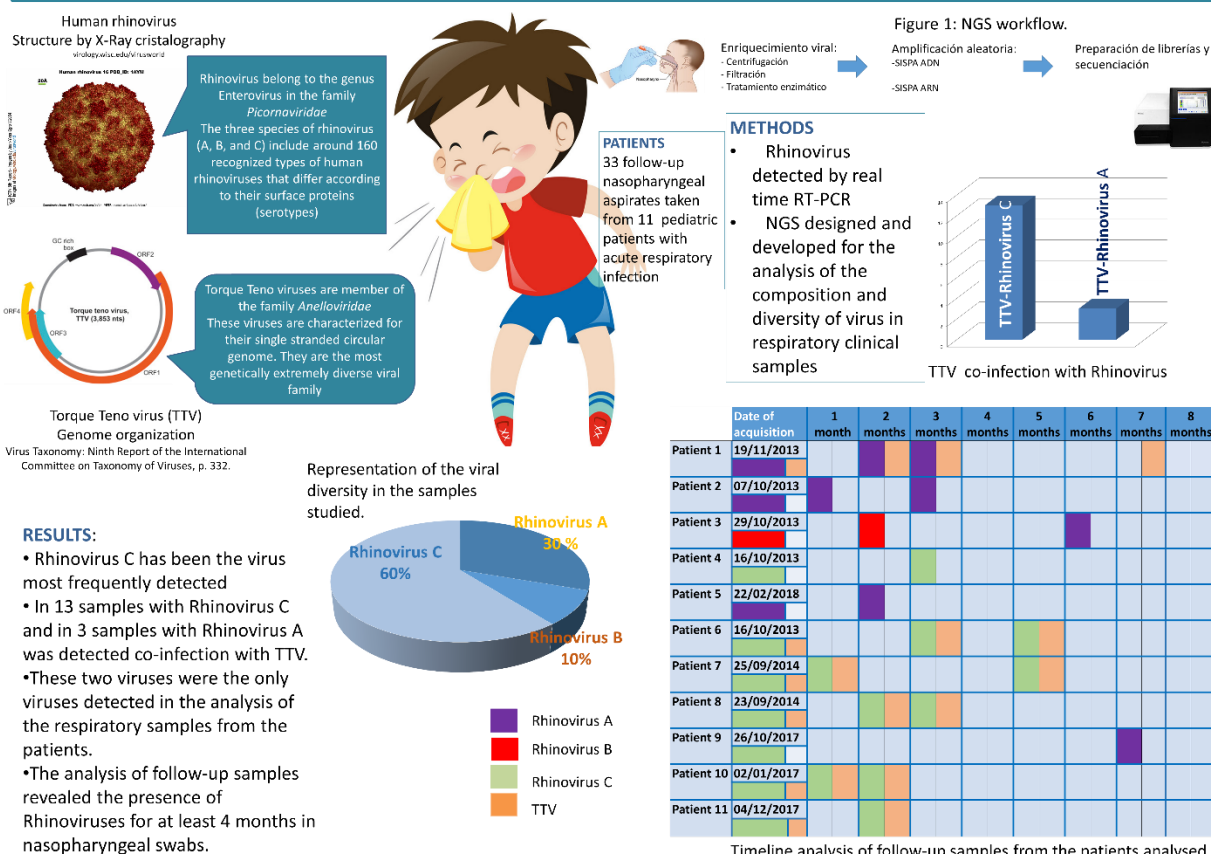
In summary, the data shown statistically different BtAdV groups present in different taxonomic bat species as hosts

# Rhinovirus and Torque Teno virus Association or coincidence?

María Iglesias-Caballero<sup>1</sup>, Unai Perez-Sautu<sup>1,2</sup>, Michael Wiley<sup>2</sup>, Francisco Pozo<sup>1</sup>, María Luz García-García<sup>3</sup>, Cristina Calvo<sup>3</sup>, Gustavo Palacios<sup>2</sup>, Inmaculada Casas<sup>1</sup>.  
1) Respiratory Virus and Influenza Unit, National Center for Microbiology, Instituto de Salud Carlos III (ISCIII), Majadahonda, Madrid, Spain; 2) Center for Genome Sciences, United States Army Medical Research Institute of Infectious Diseases (USAMRIID), Fort Detrick, Frederick, Maryland, USA; 3) Severo Ochoa University Hospital, Leganés, Madrid, Spain.

Respiratory tract infections remain responsible for nearly 3 million annual deaths predominantly in children and the elderly. A large proportion of these infections have viral etiology.

- ➔ Rhinovirus have been reported as the most frequent virus causing infection among hospitalized patients with acute respiratory illness.
- ➔ Torque Teno virus have a strong correlation between virus load and airflow limitation, suggesting might be a contributing factor in lung impairment



## CONCLUSIONS

Rhinovirus C contributes more to recurrent wheezing and exacerbations compared with Rhinovirus A and B. Rhinovirus A is more represented in previous cohort studies of pediatric patients. The largest representation of RVC in comparison with RVA in this work might be related to the severity of the patients studied. In spite of TTV role as virome resident, the viral load of TTV is connected with inflammatory processes and immune suppression. A combination of these two viruses could explain long periods of viral excretion acting as an enhancer of the infection and the upper and lower airway inflammation. The impact of this association with inflammatory diseases as asthma and if exist association among them remains to be elucidated. Partners in crime or just respiratory tract neighbors?

**Oral presentation: Novel circular rep-encoding single-stranded DNA viruses in human respiratory specimens.**

Circular Rep-encoding Single-Stranded DNA viruses, CRESS-DNA viruses, are a diverse group growing faster because of breakthroughs in the metagenomics field.

CRESS-DNA viruses were reported in unexplained cases of human encephalitis and diarrhea. Novel CRESS-DNA virus were also reported in human respiratory samples. Using high-throughput sequencing we detected a novel CRESS-DNA virus in different respiratory specimens of a patient with severe pneumonia (Index patient). Additional search was developed to investigate the detection of this Novel Respiratory CRESS-DNA virus in respiratory samples from patients with and without respiratory symptoms.

A total of 729 respiratory specimens were taken from patients over 16 years old (Group 1: 368) and pediatric patients (Group 2: 361). Specimens were extracted using manual alcohol precipitation. Three additional multiplex real-time RT-PCR assays were used for the detection of common respiratory virus.

Based on sequences, phylogenetically related with Circovirus, of the Novel Respiratory CRESS-DNA virus obtained by high-throughput sequencing of the Index patient specimens, a specific PCR was designed in the conserved REP region. Amplified products were visualized by agarose gel (381 nt) and sequenced.

Positive Novel Respiratory CRESS-DNA virus specimens from patients were analyzed by metagenomics. Libraries were pooled in an Illumina MySeq and the bioinformatics analysis was performed with a pathogen detection pipeline.




A total of 31 specimens (31/729, 4,2%) was positive to the Novel Respiratory CRESS-DNA virus, corresponding 29 from Group 1 and 3 from Group 2. Five positive patients were hospitalized with severe pneumonia and rendered negative results for common bacterial and respiratory virus. In 5 patients the Novel Respiratory CRESS-DNA virus was detected in follow-up specimens taken in different time with and without respiratory clinical symptoms. None of the control patients without symptoms was positive.

The specific pathogenic role of these new viruses in respiratory infection has to be demonstrate. Coinfections with well-known viral respiratory virus is common

We identified a Novel Respiratory CRESS-DNA viruses present in respiratory specimens taken from patients with respiratory infection with lack of significant genetic homology with the currently known CRESS-DNA viruses.

## Article

# New Adenovirus Groups in Western Palaearctic Bats

Maria Iglesias-Caballero <sup>1</sup>, Javier Juste <sup>2</sup> , Sonia Vázquez-Morón <sup>1,3</sup>, Ana Falcon <sup>1,4</sup>, Carolina Aznar-Lopez <sup>1,3</sup>, Carlos Ibáñez <sup>2</sup> , Francisco Pozo <sup>1</sup> , Guillermo Ruiz <sup>1</sup>, Jose M. Berciano <sup>1</sup>, Inazio Garin <sup>5</sup>, Joxerra Aihartza <sup>5</sup>, Juan E. Echevarría <sup>1,3</sup> and Inmaculada Casas <sup>1,\*</sup>

<sup>1</sup> Centro Nacional de Microbiología, Instituto de Salud Carlos III, Carretera de Majadahonda-Pozuelo km 2. Majadahonda 28220, Madrid, Spain; miglesias@isci.es (M.I.-C.); svazquez@isci.es (S.V.-M.); afalcon@cnb.csic.es (A.F.); carolinaaznarlopez@hotmail.com (C.A.-L.); pacopozo@isci.es (F.P.); guillermo.ruiz@salud.madrid.org (G.R.); jmberciano@isci.es (J.M.B.); jeecheva@isci.es (J.E.E.)

<sup>2</sup> Estación Biológica de Doñana, CSIC, Avda Américo Vespucio 16, 41092 Seville, Spain; juste@ebd.csic.es (J.J.); ibanez@ebd.csic.es (C.I.)

<sup>3</sup> Centro de Investigación Biomédica Epidemiología y Salud Pública, CIBERESP, 28029 Madrid, Spain

<sup>4</sup> Consorcio Centro de Investigación Biomédica en Red (CIBER), 28029 Madrid, Spain

<sup>5</sup> Department of Zoology and Animal Cell Biology, University of the Basque Country (UPV/EHU), Leioa 48940, Basque Country, Spain; inazio.garin@ehu.eus (I.G.); joxerra.aihartza@ehu.eus (J.A.)

\* Correspondence: icasas@isci.es

Received: 27 July 2018; Accepted: 18 August 2018; Published: 20 August 2018



**Abstract:** In the context of long-term screening for viruses on Western Palaearctic bats, we tested for the presence of adenovirus 1392 oropharyngeal swabs and 325 stool samples taken from 27 bat species. Adenoviruses were detected in 12 species of the Vespertilionidae and the Rhinolophidae families. Fifty positive respiratory and 26 positive stool samples were studied. Phylogenetic analyses of partial hexon protein and partial DNA-dependent DNA polymerase genes indicate that all these bat adenoviruses belong to the genus *Mastadenovirus* but without constituting a monophyletic cluster. According to genetic identities, the new groups are distinct to the previously described *Bat mastadenovirus A* and *B* species and contribute with potentially new members. Our data support that diversity of bat mastadenovirus is host-dependent and increase the knowledge of potentially pathogenic virus from bats. Due to the active role of bats as viral reservoirs, the characterization of these viruses is relevant for Public Health.

**Keywords:** Adenovirus; Western Palaearctic Bats; Phylogenetic analysis; Spain

## 1. Introduction

Bats are the second largest order of mammals, comprising more than 1200 different species [1]. Their high vagility and the organization typically in social groups predispose them to infection and viral dissemination [2]. Extensive surveys have shown their susceptibility to host a wide range of viruses and the possibility to be a source of emerging infectious in humans [3]. The Order Chiroptera plays a role as a reservoir for many significant virus families such as *Rhabdoviridae*, *Coronaviridae*, *Herpesviridae*, *Filoviridae*, *Reoviridae*, *Paramyxoviridae* and *Astroviridae*, among others. Several studies have shown bats to be a reservoir of Adenoviruses [4–7].

Adenoviruses (AdVs) are subdivided in five genera, *Mastadenovirus* (mammals), *Aviadenovirus* (birds), *Atadenovirus* (mammals, birds and reptiles), *Siadenovirus* (birds and amphibians) and *Ichtadenovirus* (fish) [8]. In 2008, the first AdV from a bat, BtAdV1-FBV1, was isolated during attempts to establish a specific cell line from a Ryukyu flying fox (*Pteropus dasymallus yayeyamae*), in Japan [9]. Following a screening of 55 German free-ranging bats, family Vespertilionidae, a second, BtAdV-2



strain PPV1, was identified in 3 common pipistrelles (*Pipistrellus pipistrellus*) [10], being the first AdV isolated from a microchiropteran bat and the second fully sequenced genome [11]. The first fully sequenced AdV genome from a bat was the BtAdV-3 strain TJM from a Rickett's big-footed bat (*Myotis ricketti*) [12]. According to ICTV, BtAdV-3 strain TJM and BtAdV-2 strain PPV1 were renamed as *Bat mastadenovirus A* and *B*. Several other studies have shown a large genetic viral diversity in bats from Brazil [13], Japan [9], Germany [4,10,11,14], China [12,15], Hungary [5,14], Ghana [16], Zambia [17], Kenya [7], South Africa [18] and USA [19].

As a consequence of these studies, the following viruses are candidates for the ICTV to be novel bat species in the future: *Bat mastadenovirus C* (*Rhinolophus sinicus* WIV9, KT698853) [6], *Bat mastadenovirus D* (*Miniopterus schreibersi* WIV12, KT698856) [15], *Bat mastadenovirus E* (*M. schreibersi* WIV13, KT698852) [15], *Bat mastadenovirus F* (*Rousettus leschenaultii* WIV17, KX961095) [15] and *Bat mastadenovirus G* (*Corynorhinus rafinesquii* 250-A, KX871230) [19].

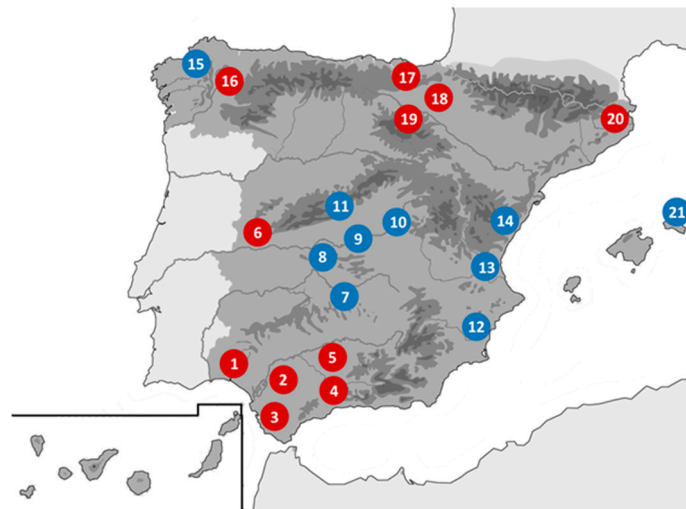
In Spain, rabies surveillance has become an important issue due to its geographic position between Africa and Europe [20], particularly bats with expected genetic flow between the South of Spain and the North of Morocco such as *Eptesicus isabellinus* [21]. Several studies have confirmed both Iberian species of *Eptesicus* as rabies vectors [22,23] including the detection of the new Lleida bat lyssavirus [24]. Other studies have described new viruses, such as a novel Lloviu filovirus detected in dead *Miniopterus schreibersi* in the North of Spain [25], 14 coronavirus distributed in new groups including two betacoronavirus related with the MERS-CoV group [26], 42 potentially novel betaherpesvirus and 10 potentially novel rhabdovirus from the families Vespertilionidae, Miniopteridae, Rhinolophidae, Molossididae and Pteropodidae in the South and North of Spain [27,28]. These studies have increased the knowledge of new viruses and their potential as human pathogens. Human AdVs cause a wide range of clinical syndromes and are being increasingly recognized in cases of severe or fatal pneumonia, haemorrhagic cystitis, hepatitis, or disseminated disease in paediatric bone marrow transplant recipients. Due to the active role of bats as viral reservoirs, this knowledge is an important part of the Public Health surveillance. Our study aimed to investigate the AdVs circulating in bats to describe their phylogenetic relationship by analysing two distinct informative partial genes.

## 2. Materials and Methods

### 2.1. Origin of Samples and Preparation

During 2004 to 2016, in the context of rabies surveillance, a screening for other different viruses was performed according to the General Research Program protocol of the Spanish Government (specific projects SAF2006-12784-C02/01-02, SAF2009-09172 and SAF2013-47194-P, approved on 10 January 2006, 20 November 2009 and 3 December 2013, respectively). Bats were captured and sampled in several campaigns across the Iberian Peninsula (Figure 1). Sampling methods followed the regulations and ethical procedures of the Spanish Bat Society (SECEMU). After being captured, each animal was identified, sexed, measured and weighed. For identification of cryptic species complexes, a wing-punch sample was collected for analysis of a cytochrome-b gene fragment [21]. For virological studies, oropharyngeal swabs (OPS) and stool samples (SS) were collected and homogenized in 1 mL of lysis buffer. After being studied and sampled, bats were released at the same location. Samples were sent to the Rabies National Reference Laboratory, aliquoted and stored at  $-80^{\circ}\text{C}$  until tested. Total nucleic acids were extracted from aliquots of 200  $\mu\text{L}$ -buffered suspension and pellets were diluted in 50  $\mu\text{L}$  of water [29].





**Figure 1.** Geographical distribution of Bat capture locations in Spain. South of Spain: 1. Huelva, 2. Seville, 3. Cádiz, 4. Málaga, 5. Córdoba. Centre of Spain: 6. Cáceres, 7. Ciudad Real, 8. Toledo, 9. Madrid, 10. Guadalajara, 11. Segovia, 12. Alicante, 13. Valencia, 14. Castellón. North of Spain: 15. A Coruña, 16. Lugo, 17. Biscay, 18. Navarra, 19. La Rioja, 20. Girona. Balearic Islands: 21. Menorca. Red circles are locations with AdV positive samples.

## 2.2. Adenovirus Detection by Generic PCR Methods

Two independent generic PCR assays were used for the detection of members of the family *Adenoviridae*. A panAdVHex nested PCR that amplified one of the seven hypervariable regions of the hexon gene and had previously been used for human AdVs genotyping was used for screening of samples [30,31]. Five  $\mu\text{L}$  of nucleic acids extracted were added to 45  $\mu\text{L}$  of reaction mixture containing 60 mM Tris-HCl (pH 8.5), 15 mM  $(\text{NH}_4)_2\text{SO}_4$ , 0.4 mM each of dNTPs (GE Healthcare, Buckinghamshire, UK), 60 pmol of each primer and 2.5 U AmpliTaq DNA Polymerase (Applied Biosystems, Branchburg, NJ, USA). Cycling conditions were: 95 °C-4 min and 40 cycles, 95 °C-30 s, 50 °C-2 min, 72 °C-30 s. For nested reactions, same reagents and cycling conditions were used. Amplified products (~768 bp) were visualized following 2% agarose gel electrophoresis. To increase the phylogenetic accuracy, a panAdVPol hemi-nested PCR assay targeting a taxonomical informative fragment of the DNA-dependent DNA polymerase gene (DNApol) was designed and used. Five  $\mu\text{L}$  of extract was added to 20  $\mu\text{L}$  of reaction mixture (LightCycler 480, Roche Diagnostics, Mannheim, Germany) and 10 pmol of the primers pol-F (5'GTIGCRAAIGAICCRTAGAGGGC 3') and pol-R (5'GTTTAYGAYATITGYGGMATGTAYGC 3'). The amplification conditions were: 95 °C-5 min, followed by 45 cycles, 95 °C-15 s, 57 °C-2 min, 68 °C-30 s. For hemi-nested reactions, 2  $\mu\text{L}$  of the previously amplified DNA and 10 pmol of the primers pol-F2 (5'AAIGAICCRTAGAGGGCCTTKGA 3') and pol-R were added to a reaction mixture containing 60 mM Tris-HCl (pH 8.5), 15 mM  $(\text{NH}_4)_2\text{SO}_4$ , 0.2 mM each of dNTPs and 1.25 U AmpliTaq DNA Polymerase. The amplification conditions were: 95 °C-5 min, followed by a two-step-cycle of 95 °C-15 s and 62 °C-2 min 45 times. Amplified products (~450 bp) were visualized following electrophoresis on a 2% agarose gel.

## 2.3. Sequence and Phylogenetic Analysis

Purified amplified products of the expected size were double-strand sequenced by Sanger chain-termination method using the BigDye Terminator v3.1 Cycle Sequencing Kit in an ABI PRISM 3700 DNA Analyzer (Applied Biosystems). The nucleotide sequences were compared with those published in GenBank database using the BLASTn algorithm (<http://blast.ncbi.nlm.nih.gov/>) to assess and identify similar AdV sequences. Two nucleotide multiple-sequence alignments from the hexon and DNApol genes, comprising a selection of available mastadenovirus sequences from the

GenBank database, were constructed using CLUSTAL X (v.2.0; <http://www.clustal.org/>). Phylogenetic analysis was performed with MEGA 5.2 software (<http://www.megasoftware.net>) and were based on a Neighbour-Joining criterion using a Tamura 3 and Kimura 2-parameter models for the hexon and DNAPol genes respectively, selected by Modeltest software [32]. Pairwise distance comparison between the predicted DNAPol amino acid sequences of Iberian bat AdVs and *Bat mastadenovirus A* and *B* was calculated using MEGA 5.2 software. Names for the putative new bat AdVs were assigned using the bat host species abbreviation and the identification ring number.

### 3. Results

Bat species studied, year of capture, type of sample and the corresponding GenBank accession numbers for the Iberian bat AdV sequences are listed in Table 1.

**Table 1.** Bat species studied, AdV positive results, year of capture and GenBank accession numbers.

<sup>1</sup> Abb., Bat species abbreviations, <sup>2</sup> OPS, Oropharyngeal swabs, <sup>3</sup> SS, Stool samples, <sup>4</sup> Capture Year,

<sup>5</sup> GenBank Accession number for hexon sequences, <sup>6</sup> GenBank Accession number for DNA polymerase sequences. N/A: Not available.

Iberian Bat Species		Abb <sup>1</sup>	OPS <sup>2</sup>	SS <sup>3</sup>	Year <sup>4</sup>	Hexon Sequences <sup>5</sup>	DNA-Pol Sequences <sup>6</sup>
Family	Name						
Vespertilionidae	<i>Barbastella barbastellus</i>	Bba	0/38	0/4	07,08	N/A	N/A
	<i>Eptesicus isabellinus</i>	Eis	0	0/8	04,07	N/A	N/A
	<i>Eptesicus serotinus</i>	Ese	0	0/14	03,07	N/A	N/A
	<i>Hypsugo savii</i>	Hsa	0/31	3/26	07	HM856338,41,42	JX065121, 22, MG208122
	<i>Myotis alcathoe</i>	Mal	0	0/1	07	N/A	N/A
	<i>Myotis bechsteinii</i>	Mbe	1/18	0/2	07	MF540611	N/A
	<i>Myotis blythii</i>	Mbl	0/29	0	04	N/A	N/A
	<i>Myotis capaccinii</i>	Mca	0/15	0	04,07	N/A	N/A
	<i>Myotis daubentonii</i>	Mda	0/63	0/41	04,07	N/A	N/A
	<i>Myotis emarginatus</i>	Mem	3/56	0	08	MF540608-10	N/A
	<i>Myotis escalerae</i>	Mes	0/13	0	04,07	N/A	N/A
	<i>Myotis myotis</i>	Mmy	1/79	0/1	04,07	HM856353	N/A
	<i>Myotis mystacinus</i>	Mmt	0/2	0/8	07	N/A	N/A
	<i>Myotis nattereri</i>	Mna	0/36	0/3	07	N/A	N/A
	<i>Nyctalus noctula</i>	Nno	3/122	0	07	MF540597-99	N/A
	<i>Nyctalus lasiopterus</i>	Nlas	10/139	6/40	07	HM856327-34,39-40,43,45-47, 50, MG132211	JX065117-20, 23,25-26,28
	<i>Nyctalus leisleri</i>	Nle	1/19	3/26	07	HM856344,48, 51-52	JX065124,27,29
	<i>Pipistrellus kuhlii</i>	Pku	12/350	2/4	07,16	MF540577-85,87,89	MF404970-73,75,86
	<i>Pipistrellus pipistrellus</i>	Ppi	0/29	0/4	07,16	HM856349	N/A
	<i>Pipistrellus pygmaeus</i>	Ppy	6/36	11/120	07,16	MF540575-76, 86,88,90-96	MF404968-69,74, 76-79, 80-85,87-89
	<i>Plecotus auritus</i>	Pau	0/11	0/8	04,07	N/A	N/A
	<i>Plecotus austriacus</i>	Pas	0/10	0/6	04,07	N/A	N/A
Miniopteridae	<i>Miniopterus schreibersii</i>	Msc	0/152	0/2	04,07, 16	N/A	N/A
Rhinolophidae	<i>Rhinolophus euryale</i>	Reu	6/49	0	04,07, 08	MF540600-02,12-13 HM856335	N/A
	<i>Rhinolophus ferrumequinum</i>	Rfe	7/90	1/3	04,07	MF540603-07,14 HM856336-37	N/A
	<i>Rhinolophus hipposideros</i>	Rhi	0/4	0/4	07,08	N/A	N/A
	<i>Rhinolophus mehelyi</i>	Rme	0/1	0	07,08	N/A	N/A
Total	27/32		50/1392	26/325		69 49OPS + 20SS	35 14OPS + 21SS

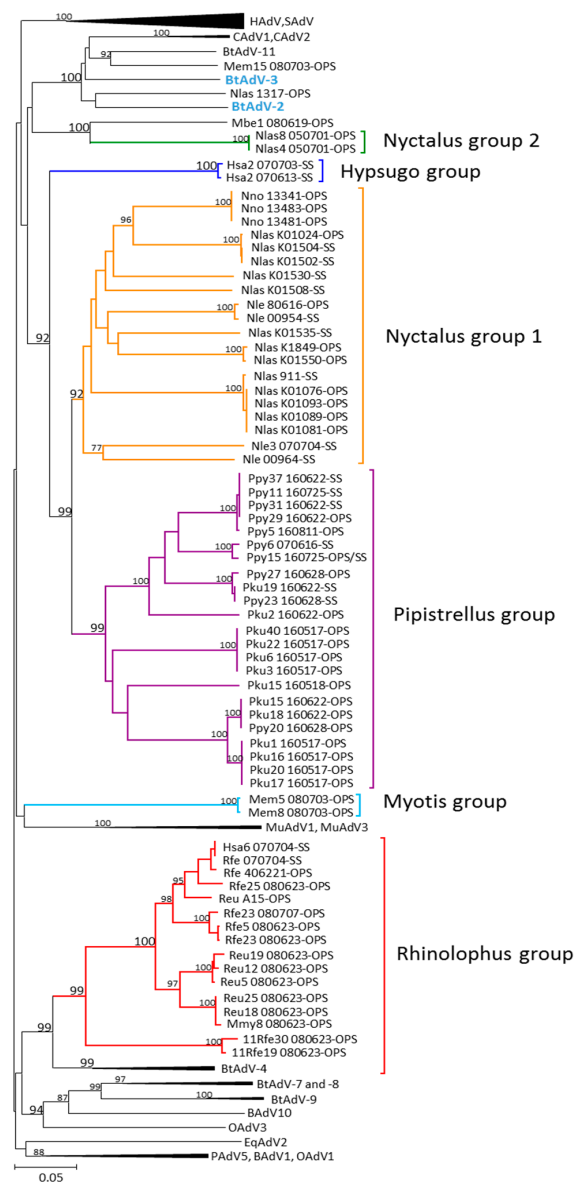
We screened a total of 1717 samples, 1392 OPS and 325 SS, representing 27 out of the 32 European bat species (<http://secemu.org>), belonging to the families Vespertilionidae (22 spp), Miniopteridae (1 sp) and Rhinolophidae (4 spp). AdV DNA was detected in 50 OPS (3.6%) and in 26 SS (8.3%). Seventy individual bats had detectable levels of AdV DNA with three of these being positive in both OPS and SS. Successful amplification of the partial AdV hexon gene was obtained in 69 samples, (49 OPS and 20 SS) and for the partial DNAPol gene in 35 samples (14 OPS and 21 SS). All amplified products were confirmed by sequencing and individual sequences were deposited in the GenBank

database (Table 1). In 29 bats both partial genes were studied. In 41 bats only the hexon sequence were obtained. Finally, in 6 bats only the DNAPol was studied.

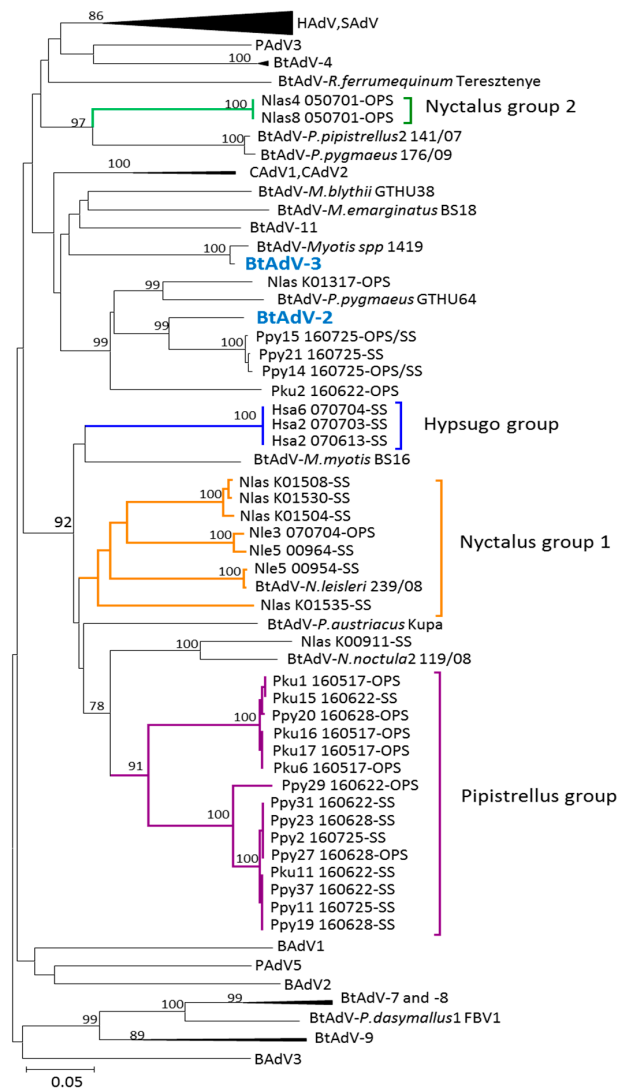
The Andalusia region in the south of Spain, had the greater distribution of AdV in bats including several genera of the families Vespertilionidae, (*Pipistrellus*, *Myotis* and *Nyctalus*) and the Rhinolophidae (*Rhinolophus*). The majority of positive bats belonging to the *Pipistrellus* genus were sampled in Andalusia. All 59 bat AdVs found in the *Rhinolophus* genus also came from Andalusian bats while no positives were detected in 78 bats sampled in the Basque Country (North). The three *Nyctalus* species (*N. noctula*, *N. lasiopterus* and *N. leisleri*, 23 bats) and two of the three *Pipistrellus* (*P. kuhlii* and *P. pygmaeus*, 28 bats) contributed the most to the list of positives detected in OPS and SS. Two out of four species of the *Rhinolophus* genus (*R. euryale* and *R. ferrumequinum*, 14 bats) had detectable levels of AdV DNA present in the OPS samples only.

### 3.1. Phylogenetic Analysis of Bat AdV Sequences

Our sequences from the partial AdV hexon and DNAPol genes, Figures 2 and 3 respectively, were included within the genus *Mastadenovirus*. High bootstrap values supported clusters which differentiate the bat mastadenovirus from the families Rhinolophidae and Vespertilionidae. Similar clustering in the phylogenetic trees using the partial hexon and DNAPol genes were observed when compared to the complete genomes sequences with high bootstrap values obtained.



**Figure 2.** Phylogenetic tree based on the analysis of the hexon partial gene. Trees were estimated with MEGA 5.2 software by using the neighbour-joining method on Tamura 3 parameters model. A bootstrap test was replicated for 5000 times. Numbers represent percentage bootstrap support. GenBank accession numbers for the sequences included in the tree are as follows: BtAdV-3 (strain TJM, GU226970, *Bat mastadenovirus A*), BtAdV-2 (strain PPV1, JN252129, *Bat mastadenovirus B*), human AdVs: type 1 (AF534906), type 2 (J01917), type 3 (DQ086466), type 4 (AY594254), D8 strain Ger/Berlin/04\_2003 (KT862545), type 9 (AJ854486), type 12 (X73487), type 14 (FJ841902), type 16 (X74662), type 21 (KF528688), type 24 (JN226751), type 27 (JN226753), type 42 (JN226761), type 45 (JN226764), simian AdVs: type 1 (AY771780), type 4 (KP853121), ovine AdVs: type 1 (DQ630754), type 3 strain (DQ630756), porcine AdV 5 (AF289262), murine AdVs: type 1 (M81889), type 3 (EU835513), bovine AdVs: type 1 (DQ630761), type 10 (AF282774), canine AdVs: type 1 (KX545420), type 2 (U77082), equine AdV type 2 (L80007), bat mastadenovirus: BtAdV-7 (strain WIV12, KT698856, *Bat mastadenovirus D*), BtAdV-8 (strain WIV13, KT698852, *Bat mastadenovirus E*), BtAdV-9 (strain WIV17, KX961095, *Bat mastadenovirus F*), *Rousettus leschenaultii* WIV18 (NC\_035072), BtAdV-4 (strain WIV9, KT698853, *Bat mastadenovirus C*), *Rhinolophus sinicus* WIV10 (NC\_029899), *R. sinicus* WIV11 (NC\_029902), BtAdV-11 (strain 250-A, KX871230, *Bat mastadenovirus G*).



**Figure 3.** Phylogenetic tree based on the analysis of the DNA-dependent DNA polymerase partial gene. Trees were estimated with MEGA 5.2 software by using the neighbour-joining method on Kimura 2 parameters model. A bootstrap test was replicated for 5000 times. Numbers represent percentage bootstrap support. GenBank accession numbers for the sequences included in the tree are as follows: BtAdV-3 (strain TJM, GU226970, *Bat mastadenovirus A*), BtAdV-2 (strain PPV1, JN252129, *Bat mastadenovirus B*), human AdVs: type 1 (AF534906), type 2 (J01917), type 3 (DQ086466), type 4 (AY594254), type 5 (AY339865), type 7 (AY594256), type 6 (HQ413315), type 9 (AJ854486), type 12 (X73487), type 17 (AF108105), type 19 (JQ326209), type 26 (EF153474), type 48 (EF153473), type 53 (AB605245), simian AdVs: type 1 (AY771780), type 4 (KP853121), bovine AdVs: type 2 (AF252854), type 3 (AF061654), type 1 (AC\_000191), porcine AdVs: type 3 (AB026117), type 5 (AF289262), canine AdVs: type 1 (KX545420), type 2 (U77082), bat mastadenovirus: BtAdV-7 (strain WIV12, KT698856, *Bat mastadenovirus D*), BtAdV-8 (strain WIV13, KT698852, *Bat mastadenovirus E*), BtAdV-9 (strain WIV17, KX961095, *Bat mastadenovirus F*), *Rousettus leschenaultii* WIV18 (NC\_035072), BtAdV-4 (strain WIV9, KT698853, *Bat mastadenovirus C*), *Rhinolophus sinicus* WIV10 (NC\_029899), *R. sinicus* WIV11 (NC\_029902), BtAdV-11 (strain 250-A, KX871230, *Bat mastadenovirus G*), *Plecotus austriacus* Kupa (JN167523), *Rhinolophus ferrumequinum* Teresztenye (JN167522), *Myotis spp 1419* (GU226962), *R. leschenaultii* 1050597 (HQ529709), *Nyctalus noctula* 2 119/08 (KM043096), *Myotis emarginatus* BS18 (KM043084), *Myotis myotis* BS16 (KM043106), *Myotis blythii* GTHU38 (KM043086), *Nyctalus leisleri* 239/08 (KM043102), *Pipistrellus pygmaeus* GTHU64 (KM043090), *P. pygmaeus* 176/09 (KM043091).

### 3.2. Partial AdV Hexon Gene Sequence Analysis

Two of the 73 bat AdVs detected in OPS from a *Nyctalus lasiopterus* (HM856343) and a *Myotis emarginatus* (MF540610) slightly related with bat adenovirus 11 (species *Bat mastadenovirus G*) (KX871230) from a *Corynorhinus rafinesquii* captured in USA [19], clustered with the reference strains bat adenovirus 2 (JN252129) and 3 (GU226970), detected in a *Pipistrellus pipistrellus* from Germany [10] and in a *Myotis ricketti* from China [12], respectively.

The rest of the 73 bat AdVs clustered in six groups with significant bootstrap values, supporting potential novel groups within the genus *Mastadenovirus* based on the bat species. These new groups are host differentiated: *Pipistrellus* group (Table S1), *Nyctalus* group 1, *Nyctalus* group 2 (Table S2), *Hypsugo* group and *Myotis* group (Table S3) of the *Vespertilionidae* family and a *Rhinolophus* group (Table S4) of the *Rhinolophidae* family (Figures 2 and 3).

*Nyctalus* group 1 represented a cluster of 13 AdVs from *N. lasiopterus*, four from *N. leisleri* and three from *N. noctula*, highly associated with the *Pipistrellus* group. *Nyctalus* group two AdVs clustered apart including two AdVs from two distinct *N. lasiopterus* and one from a *Myotis bechsteinii* (MF540611). In a well-defined *Pipistrellus* group (bootstrap 99) 13 AdVs cluster from *P. kuhlii* and 9 from *P. pygmaeus*. Similarly, the well-defined *Rhinolophus* group included eight from *R. ferrumequinum* and six from *R. euryale* and two others from a *Myotis emarginatus* (MF540609) and one from *Hypsugo savii* (HM856338). This cluster was highly supported and included three bat AdVs detected in *R. sinicus* captured in China [6]. Furthermore, two distinct AdV detected from two *Hypsugo savii* bats were grouped in one independent cluster defined as *Hypsugo* group was highly related with the *Nyctalus*-group 1 and the *Pipistrellus* group. Additionally, two AdVs detected in two *Myotis emarginatus* bats constituted a new *Myotis* group (Figure 2).

### 3.3. Partial DNA-Dependent DNA Polymerase Gene Sequences

The groups defined in this gene were clearly associated by host with lower support in some nodes and less resolution compared with the hexon partial gene analysed (Figure 3).

Five AdVs detected in *Pipistrellus pygmaeus* (MF404968, MF404979, MF404969, MF404988 and MF404978) clustered together with the reference bat adenovirus 2 (JN252129) in a group which included three AdVs detected in a *Pipistrellus kuhlii* (MF404975), in a *Nyctalus lasiopterus* (JX065123) and in a *P. pygmaeus* (KM043090) captured in Hungary [14]. This group, which included 6 AdVs found in the genus *Pipistrellus*, grouped separately from the rest of our *Pipistrellus* bat AdVs.

Sequences from the genus *Nyctalus* grouped similarly with those defined in the hexon gene with the exception of a *N. lasiopterus* (JX065117) from the group 1 which was associated with an AdV detected in a *N. noctula* (KM043110) from Hungary. The *Nyctalus* group two was clustered with two different detected in a *P. Pipistrellus* (KM043096) and in a *P. pygmaeus* (KM43091) from Hungary. The *Pipistrellus* group which contained five AdVs detected in *P. kuhlii* and 12 in *P. pygmaeus* clustered together to define a similar group to that observed with the partial hexon gene analysis. In the *Hypsugo* group a *H. savii* (MG208122) clustered in the *Hypsugo* group unlike in the hexon gene analysis where this AdV clustered in the *Rhinolophus* group. No positive results were obtained in this gene with the rhinolophid bats.

Pairwise distance matrix values obtained from the partial amino acid sequence of DNAPol, supported the new groups (Table 2). According to the pairwise distance data none of the Iberian bat AdVs were related with the species *Bat mastadenovirus A*. Likewise, three AdVs detected in *Pipistrellus pygmaeus* (Ppy14\_160725, Ppy15\_160725 and Ppy21\_160725) had an amino acid pairwise distance of 12% being these viruses similar to the species *Bat mastadenovirus B*.



**Table 2.** Spanish bat mastadenoviruses classified by the amino acid distance matrix analysis based on partial DNA-dependent DNA polymerase. <sup>1</sup> Values more than 15% are potentially new species following the demarcation criteria. Abb: P: *Pipistrellus*. N: *Nyctalus*. H: *Hypsugo*.

Group of Bat AdV	Tentative Virus Name	Abb Name	% aa Pairwise Distances <sup>1</sup>	
			BtAdV-3	BtAdV-2
bat AdVs associated with BtAdV-2	Bat mastadenovirus <i>P. pygmaeus</i> 14 160725	Ppy14 160725	31.4	11.8
	Bat mastadenovirus <i>P. pygmaeus</i> 15 160725	Ppy15 160725	31.4	12.7
	Bat mastadenovirus <i>P. pygmaeus</i> 21 160725	Ppy21 160725	32	12.2
Potentially novel bat AdVs	Bat mastadenovirus <i>N. lasiopterus</i> K01317	Nlas K01317	33.8	25.5
	Bat mastadenovirus <i>N. lasiopterus</i> K01508	Nlas K01508	35	41.8
	Bat mastadenovirus <i>N. lasiopterus</i> K01530	Nlas K01530	35.7	42.5
	Bat mastadenovirus <i>N. lasiopterus</i> K01504	Nlas K01504	35.7	42.5
	Bat mastadenovirus <i>N. leisleri</i> 3 070704	Nle3 070704	32.5	42.5
	Bat mastadenovirus <i>N. leisleri</i> 00964	Nle 00964	34.5	44
	Bat mastadenovirus <i>N. leisleri</i> 5 00954	Nle5 00954	39.7	41.8
	Bat mastadenovirus <i>N. lasiopterus</i> K01535	Nlas K01535	41.8	52.1
	Bat mastadenovirus <i>N. lasiopterus</i> K00911	Nlas K00911	43.5	46.1
	Bat mastadenovirus <i>N. lasiopterus</i> 4 050701	Nlas4 050701	42.6	38.8
	Bat mastadenovirus <i>N. lasiopterus</i> 8 050701	Nlas8 050701	42.6	38.8
	Bat mastadenovirus <i>P. kuhlii</i> 2 160622	Pku2 160622	37.5	23.8
	Bat mastadenovirus <i>P. kuhlii</i> 1 160517	Pku1 160517	39.4	39
	Bat mastadenovirus <i>P. kuhlii</i> 15 160622	Pku15 160622	39.4	39
	Bat mastadenovirus <i>P. pygmaeus</i> 20 160628	Ppy20 160628	38.7	39
	Bat mastadenovirus <i>P. kuhlii</i> 16 160517	Pku16 160517	38.7	38.3
	Bat mastadenovirus <i>P. kuhlii</i> 17 160517	Pku17 160517	38.7	38.3
	Bat mastadenovirus <i>P. kuhlii</i> 6 160517	Pku6 160517	38.7	38.3
	Bat mastadenovirus <i>P. pygmaeus</i> 29 160622	Ppy29 160622	44.8	36.5
	Bat mastadenovirus <i>P. pygmaeus</i> 31 160622	Ppy31 160622	41.2	41.6
	Bat mastadenovirus <i>P. pygmaeus</i> 23 160628	Ppy23 160628	41.2	41.6
	Bat mastadenovirus <i>P. pygmaeus</i> 2 160725	Ppy2 160725	41.2	41.6
	Bat mastadenovirus <i>P. pygmaeus</i> 27 160628	Ppy27 160628	41.2	41.6
	Bat mastadenovirus <i>P. kuhlii</i> 11 160622	Pku11 160622	42	40.9
	Bat mastadenovirus <i>P. pygmaeus</i> 37 160622	Ppy37 160622	42	40.9
	Bat mastadenovirus <i>P. pygmaeus</i> 11 160725	Ppy11 160725	42	40.9
	Bat mastadenovirus <i>P. pygmaeus</i> 19 160628	Ppy19 160628	42	40.9
	Bat mastadenovirus <i>H. savii</i> 6 070704	Hsa6 070704	41	41.7
	Bat mastadenovirus <i>H. savii</i> 2 070613	Hsa2 070613	41	41.7
	Bat mastadenovirus <i>H. savii</i> 2 070703	Hsa2 070703	41	41.7

#### 4. Discussion

In this work, we describe the detection and the phylogenetic relationships among potentially new bat mastadenoviruses and known AdVs from bats using two different partial genes. Our study shows, for the first time, their diversity in bats captured in the South of Europe and particularly in Spain a region of crucial importance for its strategic geographical placement, as a corridor between Africa and Europe.

Previous studies have shown a high diversity of AdVs found in bat species analysed across Europe, Asia and Africa [9,10,13,15,17,19]. In this study of AdV in bats, 27 out of the 32 Iberian bat species were examined obtaining positive results in 12 species from 6 bat genera. In Centre of Europe, Hungary and Germany, have also found positive results for AdVs in 9 of these 12 species [14]. With the aim of having a broad representation of the AdV diversity in the Iberian bats, a total of 1717 biological samples were analysed representing the largest AdV screening of bats for adenovirus. These bats were captured within Spain in a variety of habitats, from the Pyrenees and Cantabrian mountain ranges in the North to the Mediterranean South, considered as natural border with Africa and including several bat species with possible gene flow across the Gibraltar Strait [33].

The percentage of AdV positive bats was 3.6% in OPS and 8.3% in SS over the 18.6% in German samples and the 9.9% in Hungary [14]. These marked differences could be explained by the health of the bats and/or the use of different type of biological samples, from the homogenised internal organ tissues taken in dead or injured bats in the German study to healthy bats and guano samples in roosting places in the Hungarian. Positive bat AdV percentage similarity between our study and

the Hungarian could be explained by the type of samples studied (OPS and SS). It is noteworthy the absence of AdVs in some bats such as the bent-winged *Miniopterus schreibersii*, despite the large number of individuals of this species screened. Similar negative results were found in Germany and Hungary [14]. Most of the AdV positive bats were found within the diverse bat family Vespertilionidae and particularly within the tribe *Pipistrellini* (*Pipistrellus* and *Nyctalus*), whereas they were absent from another bat tribe *Plecotinii* (*Barbastella* and *Plecotus*). Within the subfamily *Myotinae*, bats were found to be positive in several species. Except for two *Myotis emarginatus* (MF540608, MF540609), that could represent a group based on clustering according to the analysis of the hexon partial gene, the rest of *Myotis* were sparsely along the phylogenetic trees without making any monophyletic cluster. Interestingly, AdVs were not found in some *Myotis*, *M. daubentonii*, despite this species being well represented in the screening ( $n = 60$  and  $n = 41$  for OPS and SS, respectively).

Previous studies mostly focused on the analysis of guano and internal tissues [9,10,13,14,17]. The analysis of OPS for the screening of AdV is a novel aspect of this study and has allowed the AdV detection in the upper respiratory tract of bats and highlighted a possible faecal-oral transmission route with the same AdV identified in OPS and SS samples from two *P. pygmaeus* bats (Ppy15\_160725 and Ppy14\_160725). The phylogenetic reconstructions identify, in both type of samples, AdVs highly related in different groups of bats, supporting this possible oral-faecal transmission. An important reason for the study of OPS in bats is the fact that many human AdV serotypes have not a specific well identified cellular receptor and given that replicate poorly in animals [34], the understanding of factors that define tropism and transmission during a natural infection increase the knowledge of AdV infections. Notably, in bats a possible faecal-oral transmission route is an interesting issue to explore considering bat as emerging and re-emerging infectious diseases vectors.

Previous authors have published new bat mastadenovirus mostly based on the phylogenetic analysis of a short and informative fragment of the DNAPol gene [10–12,14]. This is a well conserved gene involved in viral transcription [35]. Despite its extensive use in phylogenetic analysis of new human and animal AdVs, the resolution of the phylogenetic reconstruction based on it is limited (less than 100 amino acids). The PCR presented in this study amplified ~450 bp, offering the possibility to increase the resolution of the phylogenetic tree. However, with the aim to compare our sequences with the previously published from the Central Europe [14] and the reference sequences available in the GenBank database, the length was reduced to 277 bp. Currently, ICTV has accepted two bat AdVs species, *Bat mastadenovirus A* [12] and *Bat mastadenovirus B* [13]. According to the taxonomic criteria [8] and based on the distance matrix analysis, the bat mastadenoviruses identified in our study represent potentially new species in the genus *Mastadenovirus* and very divergent from the ICTV references, even with the potentially novel species proposed, with the exception of three detected in *P. pygmaeus* (Ppy15\_160725, Ppy21\_160725, Ppy14\_160725). Moreover, one *P. kuhlii* (Pku2\_160622) and one *N. lasiopterus* (Nlas\_K01317) were associated with the *Bat mastadenovirus B* although there was greater than 15% difference suggesting new bat AdV species. It is remarkable that bat AdVs obtained from the species *P. kuhlii* and *P. pygmaeus* clustered together in two well supported groups indicating host specificity even at the species level.

In this work, the identification of new bat AdVs is further supported by the results obtained using the hexon gene, a more variable protein [11,12,16,19] which contains seven hypervariable regions identified as viral epitopes [36]. Nucleic acids variation define the different human serotypes [37]. Our generic PCR in the hexon gene was designed in the hypervariable region 7 and the analysis of the sequences obtained were in concordance with the genotype and serotype in human AdVs [30].

The evolutionary relationships based on the two partial genes are presented separately since they provide different information according to their different mutation rates. Both genes agree in the main structure of their tree topologies and clusters and both provide support for a presumably new Iberian bat mastadenoviruses clustering and distinguishing between the families Vespertilionidae and Rhinolophidae in the phylogenies. Most of the available AdVs in the GenBank database grouped within the three monophyletic groups corresponding to their host genera *Pipistrellus*, *Nyctalus* and



*Rhinolophus*. This relationship is also supported by the phylogenetic analysis of the DNAPol gene in which the AdV detected in a *N. leisleri* (Nle\_00954) clusters with a bat AdV detected in a *N. leisleri* sampled in Hungary [14]. In our sampling, more basal relationships among the main bat hosts were more difficult to be affiliated with their hosts due to the lack of representation of important bat groups such as *Scotophilinii*, *Nycticeinii* and *Plecotinii* within the family Vespertilionidae. These host-pathogen relationships were clearly observed with herpesviruses [27] but still, AdVs could represent another example of parallel evolution of DNA virus and their bat hosts. The phylogenetic analysis of partial hexon gene showed no similarities in the sequences between AdVs from bats captured in the South and the North of Spain, as it is shown in a *P. pygmaeus* (Ppy6\_070616) collected in Lugo (North) and the *P. pygmaeus* (Ppy15\_160725) collected in Seville (South).

Although most of the AdVs clustered by their bat host, some exceptions are clearly remarkable. In the hexon gene, the AdV detected in a *M. myotis* (Mmy8\_080623) clustered with the group composed of two different species of *Rhinolophus* bats. It is well known that many *Myotis* colonies share roosts with several species of the genus *Rhinolophus* and this could be the origin of the inter-specific transmission between these two bat species. Secondly, the AdV detected in a *H. savii* (Hsa6\_070704), clustered with adenovirus from the *Rhinolophus* group despite the DNAPol gene revealing a specific AdV group in three different *H. savii*. In this second example, a natural transmission seems less likely since the two species have very different life history and barely share any ecological requirement. Nevertheless, the description of recombinant viruses is a common phenomenon in human AdV [38] and could explain the different results. However, this possible recombination in bat AdVs requires a further confirmation by the complete genomic sequence. A third exception showed the AdV detected in a *M. emarginatus* (Mem15-080703) that clustered together with a *Bat mastadenovirus G* detected in a *Corynorhinus rafinesquii* bat and two others detected in *Myotis* bats from Hungary. The *C. rafinesquii* is a vespertilionid bat, the distribution of which is restricted to the Southeast of North America and Mexico [19,39]. The connection between these viruses is an intriguing given that their hosts are geographically and evolutionary distant, although it could be related to a recent colonization of North America by Palearctic *Myotis* [40].

In conclusion, based on the analysis of two different regions of genome used to study two different type of samples, the present study contributes with potentially new members from *Mastadenovirus* genus distinct from previously described reference species *Bat mastadenovirus A* and *B* [10,12]. The new AdV groups were detected in bats captured in a broad geographical region and generate data supporting that diversity of bat mastadenovirus is associated by host and the distribution of the host.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/1999-4915/10/8/443/s1>, Table S1: Tentative virus names, type of sample and localization of *Pipistrellus* group, Table S2: Tentative virus names, type of sample and localization of *Nyctalus* group, Table S3: Tentative virus names, type of sample and localization of *Hypsugo* and *Myotis* groups and Table S4: Tentative virus names, type of sample and localization of *Rhinolophus* group.

**Author Contributions:** Conceptualization, I.C., F.P., J.J., C.I. and J.E.E.; Methodology, M.I.-C. and I.C.; Validation, I.C. and M.I.-C.; Formal Analysis, M.I.-C., I.C., J.J., C.I., A.F., F.P. and G.R.; Investigation, M.I.-C., I.C., A.F. and G.R.; Resources, J.J., C.I., J.M.B., I.G. and J.A.; Data Curation, M.I.-C., S.V.-M., C.A.-L. and I.C.; Writing-Original Draft Preparation, M.I.-C. and I.C.; Writing-Review & Editing, M.I.-C., I.C. and J.J.; Supervision, I.C., J.J. and J.E.E.; Project Administration, J.E.E., S.V.-M. and J.J.; Funding Acquisition, J.E.E.

**Funding:** This project was financially supported by an agreement between the Public Health Department of the Spanish Ministry of Health and the Instituto de Salud Carlos III for the development of “Rabies Surveillance in Spain” and by projects SAF 2006-12784-C02-01, SAF 2006-12784-C02-02, SAF 2009-09172 and SAF2013-47194-P of the General Research Programme of the Spanish Ministry of Science and Education.

**Acknowledgments:** We thank the Genomics Unit of the Instituto de Salud Carlos III for the sequencing of amplified products. We thank all members of the Spanish Bat Conservation Society (SECEMU) that generously contributed with their samples to the present study.

**Conflicts of Interest:** None of the authors of this paper has a financial or personal relationship with other people or organisations that could inappropriately influence or bias the content of the paper.

## References

1. IUCN SSC Bat Specialist Group. Available online: <http://www.iucnbsg.org/> (accessed on 14 May 2018).
2. Calisher, C.H.; Childs, J.E.; Field, H.E.; Holmes, K.V.; Schountz, T. Bats: Important reservoir hosts of emerging viruses. *Clin. Microbiol. Rev.* **2006**, *19*, 531–545. [[CrossRef](#)] [[PubMed](#)]
3. Wong, S.; Lau, S.; Woo, P.; Yuen, K.Y. Bats as a continuing source of emerging infections in humans. *Rev. Med. Virol.* **2007**. [[CrossRef](#)] [[PubMed](#)]
4. Drexler, J.F.; Corman, V.M.; Wegner, T.; Tateno, A.F.; Zerbinati, R.M.; Gloza-Rausch, F.; Seebens, A.; Müller, M.A.; Drosten, C. Amplification of emerging viruses in a bat colony. *Emerg. Infect. Dis.* **2011**, *17*, 449–456. [[CrossRef](#)] [[PubMed](#)]
5. Jánoska, M.; Vidovszky, M.; Molnár, V.; Liptovszky, M.; Harrach, B.; Benkő, M. Novel adenoviruses and herpesviruses detected in bats. *Vet. J.* **2011**, *189*, 118–121. [[CrossRef](#)] [[PubMed](#)]
6. Tan, B.; Yang, X.-L.; Ge, X.-Y.; Peng, C.; Zhang, Y.-Z.; Zhang, L.-B.; Shi, Z.-L. Novel bat adenoviruses with an extremely large E3 gene. *J. Gen. Virol.* **2016**, *97*, 1625–1635. [[CrossRef](#)] [[PubMed](#)]
7. Waruhiu, C.; Ommeh, S.; Obanda, V.; Agwanda, B.; Gakuya, F.; Ge, X.Y.; Yang, X.L.; Wu, L.J.; Zohaib, A.; Hu, B.; et al. Molecular detection of viruses in Kenyan bats and discovery of novel astroviruses, caliciviruses and rotaviruses. *Virol. Sin.* **2017**, *32*, 101–114. [[CrossRef](#)] [[PubMed](#)]
8. Harrach, B.; Benkő, M.; Both, G.; Brown, M.; Davison, A.J.; Echavarría, M.; Hess, M.; Jones, M.S.; Kajon, A.; Lehmkuhl, A.D.; et al. Family Adenoviridae. In *Virus Taxonomy: Classification and Nomenclature of Viruses. Ninth report of the International Committee of Taxonomy of Viruses*; Elsevier: San Diego, CA, USA, 2011; pp. 125–141.
9. Maeda, K.; Hondo, E.; Terakawa, J.; Kiso, Y.; Nakaichi, N.; Endoh, D.; Sakai, K.; Morikawa, S.; Mizutani, T. Isolation of novel adenovirus from fruit bat (*Pteropus dasymallus yayeyamae*). *Emerg. Infect. Dis.* **2008**, *14*, 347–349. [[CrossRef](#)] [[PubMed](#)]
10. Sonntag, M.; Mühldorfer, K.; Speck, S.; Wibbelt, G.; Kurth, A. New adenovirus in bats, Germany. *Emerg. Infect. Dis.* **2009**, *15*, 2052–2055. [[CrossRef](#)] [[PubMed](#)]
11. Kohl, C.; Vidovszky, M.; Mühldorfer, K.; Dabrowski, P.W.; Radonic, A.; Nitsche, A.; Wibbelt, G.; Kurth, A.; Harrach, B. Genome analysis of bat adenovirus 2: Indications of interspecies transmission. *J. Virol.* **2012**, *86*, 1888–1892. [[CrossRef](#)] [[PubMed](#)]
12. Li, Y.; Ge, X.; Zhang, H.; Zhou, P.; Zhu, Y.; Zhang, Y.; Yuan, J.; Wang, L.F.; Shi, Z. Host range, prevalence, and genetic diversity of adenoviruses in bats. *J. Virol.* **2010**, *84*, 3889–3897. [[CrossRef](#)] [[PubMed](#)]
13. Lima, F.E.; Cibulski, S.P.; Elesbao, F.; Carnieli Junior, P.; Batista, H.B.; Roehe, P.M.; Franco, A.C. First detection of adenovirus in the vampire bat (*Desmodus rotundus*) in Brazil. *Virus Genes* **2013**, *47*, 378–381. [[CrossRef](#)] [[PubMed](#)]
14. Vidovszky, M.; Kohl, C.; Boldogh, S.; Görföl, T.; Wibbelt, G.; Kurth, A.; Harrach, B. Random sampling of the Central European bat fauna reveals the existence of numerous hitherto unknown adenoviruses. *Acta Vet. Hung.* **2015**, *63*, 508–525. [[CrossRef](#)] [[PubMed](#)]
15. Tan, B.; Yang, X.L.; Ge, X.Y.; Peng, C.; Liu, H.Z.; Zhang, Y.Z.; Zhang, L.B.; Shi, Z.L. Novel bat adenoviruses with low G + C content shed new light on the evolution of adenoviruses. *J. Gen. Virol.* **2017**, *98*, 739–748. [[CrossRef](#)] [[PubMed](#)]
16. Baker, K.S.; Leggett, R.M.; Bexfield, N.H.; Alston, M.; Daly, G.; Todd, S.; Tachedjian, M.; Holmes, C.E.G.; Crameri, S.; Wang, L.-F.; et al. Metagenomic study of the viruses of African straw-coloured fruit bats: Detection of a chiropteran poxvirus and isolation of a novel adenovirus. *Virology* **2013**, *441*, 95–106. [[CrossRef](#)] [[PubMed](#)]
17. Ogawa, H.; Kajihara, M.; Nao, N.; Shigeno, A.; Fujikura, D.; Hang'ombe, B.M.; Mweene, A.S.; Mutemwa, A.; Squarre, D.; Yamada, M.; et al. Characterization of a Novel Bat Adenovirus Isolated from Straw-Colored Fruit Bat (*Eidolon helvum*). *Viruses* **2017**, *9*, 371. [[CrossRef](#)] [[PubMed](#)]
18. Van Vuren, P.J.; Allam, M.; Wiley, M.R.; Ismail, A.; Storm, N.; Birkhead, M.; Markotter, W.; Palacios, G.; Paweska, J.T. A novel adenovirus isolated from the Egyptian fruit bat in South Africa is closely related to recent isolates from China. *Sci. Rep.* **2018**, *8*, 9584. [[CrossRef](#)] [[PubMed](#)]

19. Hackenbrack, N.; Rogers, M.B.; Ashley, R.E.; Keel, M.K.; Kubiski, S.V.; Bryan, J.A.; Ghedin, E.; Holmes, E.C.; Hafenstein, S.L.; Allison, A.B. Evolution and Cryo-electron Microscopy Capsid Structure of a North American Bat Adenovirus and Its Relationship to Other Mastadenoviruses. *J. Virol.* **2017**, *91*. [[CrossRef](#)] [[PubMed](#)]
20. Mingo-Casas, P.; Sandoñs, V.; Vázquez-Morón, S.; Berciano, J.M.; Juste, J.; Echevarría, J.E. Rabies in Spain. A Peculiarity in Eurasia. *Ann. Virol. Res.* **2017**, *3*, 1030.
21. Juste, J.; Bilgin, R.; Muñoz, J.; Ibáñez, C. Mitochondrial DNA signatures at different spatial scales: From the effects of the Straits of Gibraltar to population structure in the meridional serotine bat (*Eptesicus isabellinus*). *Heredity* **2009**, *103*, 178–187. [[CrossRef](#)] [[PubMed](#)]
22. Vázquez-Morón, S.; Juste, J.; Ibáñez, C.; Berciano, J.M.; Echevarría, J.E. Phylogeny of European Bat Lyssavirus 1 in *Eptesicus isabellinus* Bats, Spain. *Emerg. Infect. Dis.* **2011**, *17*, 520–523. [[CrossRef](#)] [[PubMed](#)]
23. Vázquez-Morón, S.; Juste, J.; Ibáñez, C.; Ruiz-Villamor, E.; Avellón, A.; Vera, M.; Echevarría, J.E. Endemic Circulation of European Bat Lyssavirus Type 1 in Serotine Bats, Spain. *Emerg. Infect. Dis.* **2008**, *14*, 1263–1266. [[CrossRef](#)] [[PubMed](#)]
24. Ceballos, N.A.; Morón, S.V.; Berciano, J.M.; Nicolás, O.; López, C.A.; Juste, J.; Nevado, C.R.; Setién, Á.A.; Echevarría, J.E. Novel Lyssavirus in Bat, Spain. *Emerg. Infect. Dis.* **2013**, *19*, 793–795. [[CrossRef](#)] [[PubMed](#)]
25. Negredo, A.; Palacios, G.; Vázquez-Morón, S.; González, F.; Dopazo, H.; Molero, F.; Juste, J.; Quetglas, J.; Savji, N.; de la Cruz Martínez, M.; et al. Discovery of an ebolavirus-like filovirus in europe. *PLoS Pathog.* **2011**, *7*, e1002304. [[CrossRef](#)] [[PubMed](#)]
26. Falcón, A.; Vázquez-Morón, S.; Casas, I.; Aznar, C.; Ruiz, G.; Pozo, F.; Perez-Breña, P.; Juste, J.; Ibáñez, C.; Garin, I.; et al. Detection of alpha and betacoronaviruses in multiple Iberian bat species. *Arch. Virol.* **2011**, *156*, 1883–1890. [[CrossRef](#)] [[PubMed](#)]
27. Pozo, F.; Juste, J.; Vázquez-Morón, S.; Aznar-López, C.; Ibáñez, C.; Garin, I.; Aihartza, J.; Casas, I.; Tenorio, A.; Echevarría, J.E. Identification of Novel Betaherpesviruses in Iberian Bats Reveals Parallel Evolution. *PLoS ONE* **2016**, *11*, e0169153. [[CrossRef](#)] [[PubMed](#)]
28. Aznar-Lopez, C.; Vazquez-Moron, S.; Marston, D.A.; Juste, J.; Ibanez, C.; Berciano, J.M.; Salsamendi, E.; Aihartza, J.; Banyard, A.C.; McElhinney, L.; et al. Detection of rhabdovirus viral RNA in oropharyngeal swabs and ectoparasites of Spanish bats. *J. Gen. Virol.* **2013**, *94*, 69–75. [[CrossRef](#)] [[PubMed](#)]
29. Casas, I.; Powell, L.; Klapper, P.E.; Cleator, G.M. New method for the extraction of viral RNA and DNA from cerebrospinal fluid for use in the polymerase chain reaction assay. *J. Virol. Methods* **1995**, *53*, 25–36. [[CrossRef](#)]
30. Calvo, C.; García-García, M.L.; Sanchez-Dehesa, R.; Román, C.; Tabares, A.; Pozo, F.; Casas, I. Eight Year Prospective Study of Adenoviruses Infections in Hospitalized Children. Comparison with Other Respiratory Viruses. *PLoS ONE* **2015**, *10*, e0132162. [[CrossRef](#)] [[PubMed](#)]
31. Casas, I.; Avellon, A.; Mosquera, M.; Jabadó, O.; Echevarria, J.E.; Campos, R.H.; Rewers, M.; Perez-Breña, P.; Lipkin, W.I.; Palacios, G. Molecular Identification of Adenoviruses in Clinical Samples by Analyzing a Partial Hexon Genomic Region. *J. Clin. Microbiol.* **2005**, *43*, 6176–6182. [[CrossRef](#)] [[PubMed](#)]
32. Posada, D.; Crandall, K.A. MODELTEST: Testing the model of DNA substitution. *Bioinformatics* **1998**, *14*, 817–818. [[CrossRef](#)] [[PubMed](#)]
33. García-Mudarra, J.L.; Ibáñez, C.; Juste, J. The Straits of Gibraltar: Barrier or bridge to Ibero-Moroccan bat diversity? *Biol. J. Linn. Soc.* **2009**, *96*, 434–450. [[CrossRef](#)]
34. Zhang, Y.; Bergelson, J.M. Adenovirus Receptors. *J. Virol.* **2005**, *79*, 12125–12131. [[CrossRef](#)] [[PubMed](#)]
35. Choi, K.H. Viral Polymerases. *Adv. Exp. Med. Biol.* **2012**, *726*, 267–304. [[CrossRef](#)] [[PubMed](#)]
36. Roberts, D.M.; Nanda, A.; Havenga, M.J.; Abbink, P.; Lynch, D.M.; Ewald, B.A.; Liu, J.; Thorner, A.R.; Swanson, P.E.; Gorgone, D.A.; et al. Hexon-chimaeric adenovirus serotype 5 vectors circumvent pre-existing anti-vector immunity. *Nature* **2006**, *441*, 239–243. [[CrossRef](#)] [[PubMed](#)]
37. Rux, J.J.; Kuser, P.R.; Burnett, R.M. Structural and phylogenetic analysis of adenovirus hexons by use of high-resolution X-ray crystallographic, molecular modeling, and sequence-based methods. *J. Virol.* **2003**, *77*, 9553–9566. [[CrossRef](#)] [[PubMed](#)]
38. Kajon, A.E.; Dickson, L.M.; Murtagh, P.; Viale, D.; Carballal, G.; Echavarria, M. Molecular Characterization of an Adenovirus 3–16 Intertypic Recombinant Isolated in Argentina from an Infant Hospitalized with Acute Respiratory Infection. *J. Clin. Microbiol.* **2010**, *48*, 1494–1496. [[CrossRef](#)] [[PubMed](#)]

39. Arroyo-Cabrales, J.; Álvarez-Castañeda, S.T. *Corynorhinus rafinesquii*. *The IUCN Red List of Threatened Species 2017*; International Union for Conservation of Nature: Gland, Switzerland, 2017.
40. Stadelmann, B.; Lin, L.-K.; Kunz, T.H.; Ruedi, M. Molecular phylogeny of New World Myotis (Chiroptera, Vespertilionidae) inferred from mitochondrial and nuclear DNA genes. *Mol. Phylogenet. Evol.* **2007**, *43*, 32–48. [\[CrossRef\]](#) [\[PubMed\]](#)



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

## Capítulo 3

---

### VARIABILIDAD GENÉTICA Y ANTIGÉNICA DE LOS VIRUS DE LA GRIPE

Inmaculada Casas Flecha,  
María de la Montaña Iglesias Caballero,  
Francisco Pozo Sánchez



## Capítulo 3

### VARIABILIDAD GENÉTICA Y ANTIGÉNICA DE LOS VIRUS DE LA GRIPE

Inmaculada Casas Flecha,  
María de la Montaña Iglesias Caballero,  
Francisco Pozo Sánchez

#### 3.1 Introducción

En los virus de la gripe se identifican dos de los tres mecanismos de variabilidad genética descritos en los virus ARN. Por una parte, la capacidad de introducir mutaciones (deleciones, inserciones y cambios nucleotídicos), y por otra, la reorganización o el intercambio de genes dado que su genoma es segmentado. La variabilidad genética y antigénica configura partículas virales con mayor eficiencia biológica (*fitness* viral), lo que significa que existe capacidad de adaptación al hospedador o a una población de hospedadores. Los virus con mayor eficiencia biológica circulan, mayoritariamente o no, entre una población susceptible, y su estudio se basa en la comparación de sus características genéticas y antigénicas con aquellas de los virus de referencia de los que proceden. La selección natural, en el caso de los virus de la gripe, se ve incrementada por la adaptación a su hospedador, y por ello los cambios genéticos y antigénicos se producen de manera gradual.

Los virus de la gripe tipo A y tipo B están compuestos por ocho segmentos genéticos independientes de ARN de polaridad negativa que codifican, cada uno de ellos, las diferentes proteínas virales: los segmentos 1, 2 y 3 de la polimerasa básica 2 (PB2), polimerasa básica 1 (PB1) y polimerasa ácida (PA) que codifican la

tres subunidades de la ARN polimerasa dependiente de ARN (RdRP) y otras proteínas accesorias que producen muerte celular PB1-F2<sup>1</sup> y modulan la patogenicidad PA-X9<sup>2</sup>; el segmento 5 de la nucleoproteína (NP); el segmento 7 de la matriz (M), que codifica la proteína de la matriz 1 (M1) y la proteína de superficie (M2), que actúa como canal iónico; el segmento 8 de la proteína no estructural (NS), que codifica tanto la proteína 1 (NS1) relacionada con la evasión inmunitaria como la proteína nuclear de exportación (NEP o NS2), que regula la salida de los complejos de ribonucleoproteína viral del núcleo celular; y finalmente los segmentos 4 y 6, que codifican las glicoproteínas de superficie hemaglutinina (HA) y neuraminidasa (NA), encargadas del inicio de la infección y de la liberación de partículas virales.

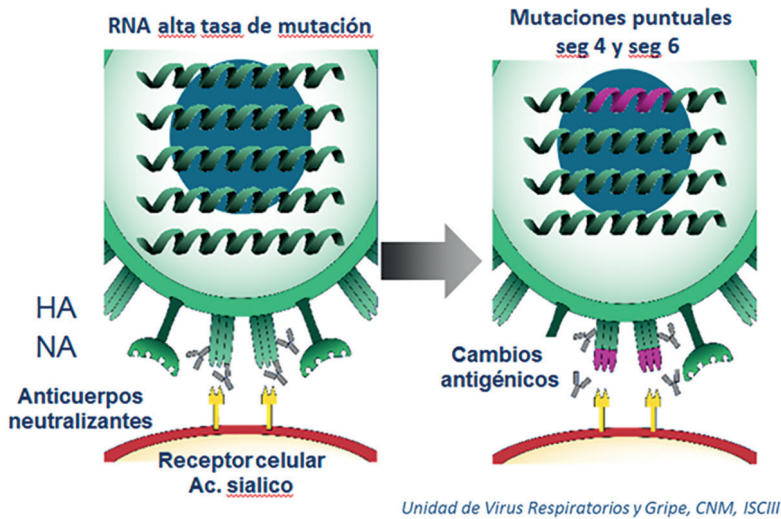
#### 3.2 Causas de la variabilidad en los virus de la gripe

Los virus de la gripe presentan dos características fundamentales por las que se les puede identificar como unos de los virus más variables que existen en la naturaleza. La primera de estas características se denomina «deriva genética» o «cambios menores» (en inglés, *drift* genético) (Figura 3.1), y la segunda se denomina «salto genético» o «cambios mayores» (en inglés, *shift* genético) (Figura 3.2). Estas



**Figura 3.1**

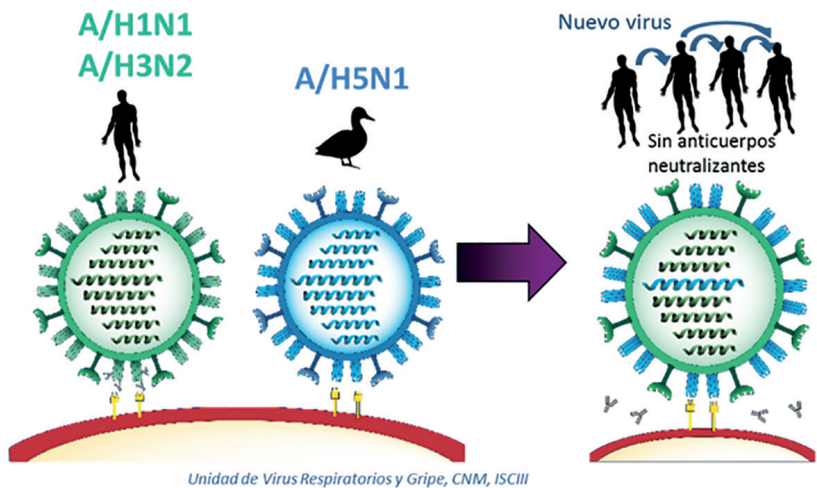
Esquema ilustrativo del mecanismo de variabilidad viral basado en los cambios genéticos (*drift* genético) y sus consecuencias en la producción de variantes antigénicas sin control por los anticuerpos neutralizantes.



**Figura 3.2**

Esquema ilustrativo de la reorganización genética como base de los cambios antigénicos mayores (*shift* antigénico).

### GRYPE A: POTENCIAL PANDÉMICO





dos características constituyen dos estrategias muy importantes para que los virus de la gripe evolucionen en función del tiempo y produzcan infecciones zoonóticas, y son básicas en los modelos de transmisión de los virus. Tanto la deriva genética como los saltos genéticos o la reorganización de segmentos dan lugar a proteínas derivadas de un virus anterior o virus con segmentos procedentes de otros virus diferentes. Aún en el siglo XXI, la deriva genética, o cambios menores, así como las reorganizaciones genéticas, o cambios mayores, son los desafíos más importantes a los que el ser humano se enfrenta al pretender establecer un control eficaz de la infección gripal mediante las vacunas y los fármacos antivirales.

### 3.2.1 Deriva genética o cambios menores (*drift* genético)

Los virus de la gripe contienen una ARN polimerasa dependiente de ARN (RdRP) incluida en el virión, debido a que su actividad no está disponible en las células del hospedador. Su función se basa en copiar los segmentos de ARN de polaridad negativa (los ocho segmentos genómicos de los virus gripales) en una cadena de ARN que funcionará como un ARN mensajero. Esta polimerasa carece de capacidad correctora, produce errores durante la copia de los segmentos genómicos en la fase de replicación viral y, como resultado, se producen fallos que dan lugar a progenies virales variantes. La RdRP viral presenta una tasa de error de  $1 \times 10^{-3}$  a  $1 \times 10^{-5}$ , lo que significa que introduce una mutación por cada 1000 a 100.000 nucleótidos copiados. A modo de ejemplo, los virus de la gripe tipo A muestran el mayor número de mutaciones acumuladas a través del tiempo, mientras que los genomas de los virus de la gripe tipo C tienen poca variabilidad. Se ha estimado que la tasa de mutación para la HA viral del subtipo A/H3 (específicamente en el dominio HA1) es de  $6,7 \times 10^{-3}$  mutaciones/sitio/año<sup>3</sup>. Ante la circulación pandémica del vi-

rus A/H1N1pdm 2009 y mediante el análisis del genoma completo de 290 aislamientos se pudo demostrar que en los primeros 4 meses de circulación el virus se diversificó en siete grupos genéticos<sup>4</sup> (Figura 3.3).

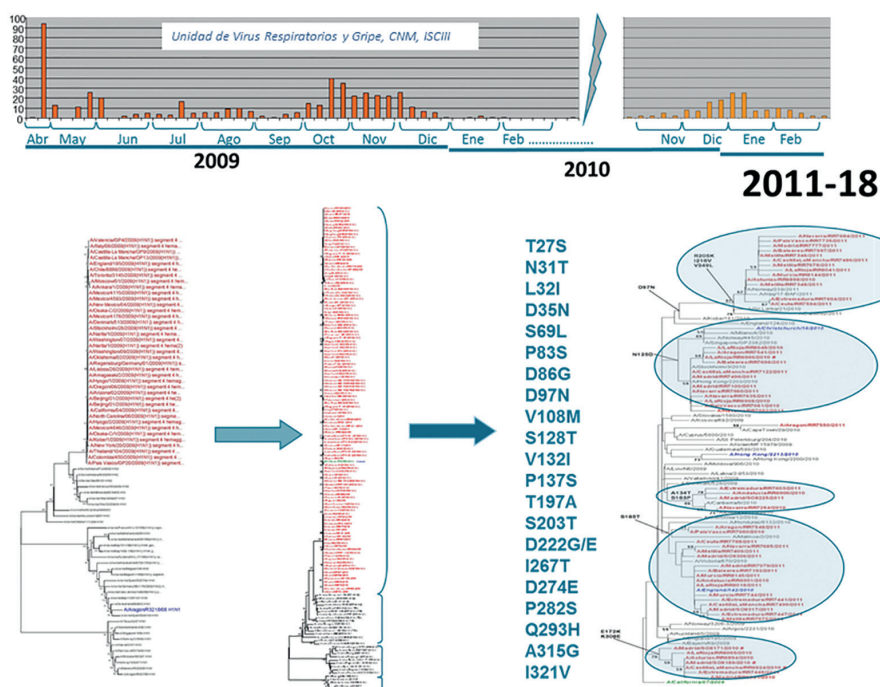
Un aspecto importante que cabe destacar es que muchas de estas mutaciones se mantienen silentes, son sustituciones sinónimas de aminoácidos en la proteína correspondiente, pero otras se traducen en aminoácidos cuyo cambio es crucial en las características de la proteína codificada (características antigénicas). En el caso concreto de la HA, se producen variantes con la capacidad de evadir los mecanismos de control que desarrolla el individuo infectado mediante la inmunidad, por lo que los fallos o errores de la RdRP durante la replicación producen variantes que permiten «escapar» del control inmunitario (Figura 3.1). Atendiendo a los subtipos de HA de los virus de la gripe A, los subtipos H4 y H13 son los de menor variabilidad, lo que sugiere que la dispersión de este tipo de virus en las aves que infectan se realiza mediante una expansión clonal. Otro ejemplo son los virus del subtipo H5, que fueron los causantes de los brotes epizooticos de los años 2014 y 2015, y que tuvieron este tipo de expansión entre las poblaciones de aves carentes de anticuerpos neutralizantes<sup>5</sup>.

Se ha demostrado, en las aves silvestres, que los segmentos 1 y 8 que codifican las proteínas PB2 y NS, que no están relacionadas con el sistema inmunitario por ser genes internos, son más diversos dependiendo del tipo de ave que infectan. Así, en las gaviotas y los patos estos dos segmentos muestran mayor variabilidad que los correspondientes segmentos procedentes de virus de otras especies de aves, y se concluye que la variabilidad depende del hospedador<sup>6</sup>.

Finalmente, se han descrito determinantes moleculares que influyen en la persisten-

**Figura 3.3**

Variabilidad genética del nuevo virus de la gripe A/H1N1pdm desde su primera detección en abril de 2009 en el Centro Nacional de Microbiología. Definición de seis grupos genéticos en un corto periodo de tiempo. Se muestran las mutaciones específicas que definen los nuevos grupos. Ninguno de estos grupos genéticos ha sido diferente antigénicamente del virus utilizado en la vacuna con la cepa A/California/07/2009.



cia de los virus de la gripe y han redefinido los modelos ecológicos de su transmisión e incluso la diversidad genética en el medio ambiente<sup>7</sup>.

### 3.2.2 Salto genético o cambios mayores (*shift* genético)

Las aves acuáticas son los reservorios naturales de los virus de la «gripe aviar», y se ha demostrado que pueden dispersarlos a escala intercontinental mediante sus migraciones. El genoma fragmentado permite la posibilidad de reorganización cuando dos virus diferentes, simultáneamente, infectan

una célula y se replican en ella, permitiendo mezclar diferentes linajes de virus entre sí<sup>8</sup> (Figura 3.2). Las reorganizaciones de virus gripales han sido la causa de la ocurrencia de pandemias en el ser humano y de diferentes brotes epidémicos en las aves domésticas de corral. Dada su importancia, los estudios de evolución de los virus aviáres y sus reorganizaciones ilustran la dinámica viral en el ecosistema que, además, está compuesto por diferentes hospedadores. Todos los segmentos, con excepción del segmento 3 que codifica la proteína PA, presentan patrones de agrupamiento por similitud genética que definen los «grupos genéticos» o «clados», que están definidos por el hospedador, con-

firmando que el flujo genético de los virus gripales en una población de aves silvestres diferentes está segregado por el tipo de ave, y además se han demostrado bajas tasas de reorganización de segmentos entre diferentes hospedadores<sup>5</sup>. Conocer la evolución viral en el ecosistema natural con una mezcla compleja de hospedadores es clave para la investigación de los virus de la gripe emergentes y del riesgo de la introducción de variantes emergentes y su circulación de una región geográfica a otra<sup>6</sup>.

La última pandemia que sufrió el ser humano se produjo por la emergencia, en 2009, de un virus de la gripe A procedente de una reorganización entre virus porcinos H1N1 pertenecientes al linaje Eurasiático (segmentos de la NA y la M) y virus porcinos H2N1 procedentes a su vez de una triple reorganización entre virus porcinos H1N1 pertenecientes al linaje clásico o americano, un virus aviar y un virus humano H2N2<sup>9</sup>. Todos los segmentos del genoma del nuevo virus, excepto el segmento PB1, tienen origen en las aves y a partir de ellas se infectaron los cerdos. Se ha demostrado que en 1918 se produjeron infecciones en cerdos por virus aviares que poseían los genes HA, NP y NS, que definen los virus del linaje porcino clásico o americano. A finales de la década de 1990 se hizo enzoótico en la cabaña americana un virus en el que se identificó un triple reagrupamiento genético<sup>10</sup>.

### 3.3 Dianas virales y metodología para el estudio de la variabilidad genética

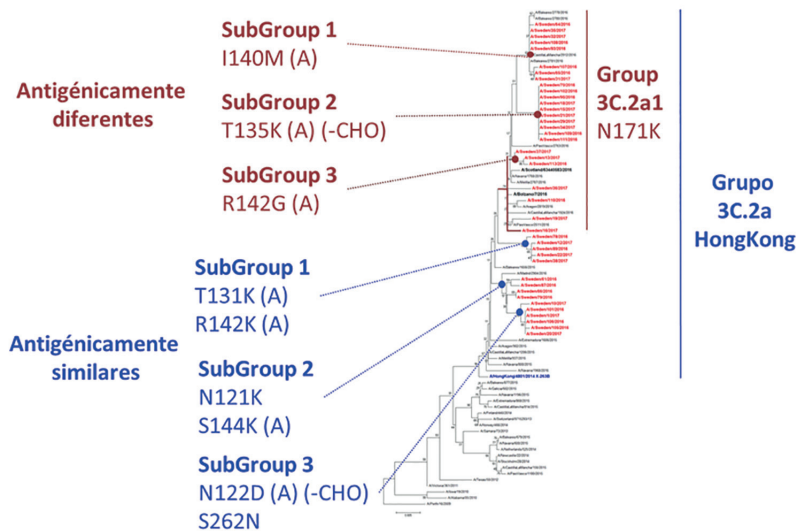
En los estudios de variabilidad aplicados a la vigilancia virológica como objetivo principal de salud pública se investigan los dos genes que codifican las dos proteínas de superficie: la HA y la NA. Las proteínas de superficie determinan las características virales más relevantes en salud pública:

- La manera en que los virus de la gripe responden a los medicamentos antivirales.
- La semejanza genética con los virus que componen la vacuna antigripal de una determinada temporada epidémica de gripe (Figuras 3.4 y 3.5).
- El potencial que pueden llegar a desarrollar para llevar a cabo saltos de barrera interespecies y de ahí el potencial zoonótico de los virus de la gripe animal (aviar, porcina, etc.).

La secuenciación de genomas es un proceso que determina el orden o la secuencia de los nucleótidos (A, U/T, C o G) en cada uno de los genes presentes en el genoma de un virus. En el caso de los virus de la gripe, la secuenciación completa de sus ocho segmentos resulta en una secuencia total de alrededor de 13.500 nucleótidos, mientras que la secuenciación parcial muestra la secuencia de algunos segmentos o genes, o incluso determinadas partes de ellos que son seleccionadas por su concordancia con las características genéticas del gen completo para su estudio. Este sistema de análisis se denomina «caracterización genética» y se realiza en los laboratorios de salud pública de todo el mundo desde los años 1980. Las secuencias genéticas obtenidas se almacenan en bases de datos públicas, como GenBank (<https://www.ncbi.nlm.nih.gov/genbank/>), y desde hace unos años también en la Global Initiative on Sharing All Influenza Data (GISAID; <https://www.gisaid.org/>). Las secuencias almacenadas permiten comparar los genes completos, o fragmentos de ellos, de los virus de la gripe que circulan durante una determinada temporada epidémica con las secuencias de los genes de los virus gripales que circularon anteriormente y con las secuencias de los virus componentes de las vacunas (virus vacunales o referencia). Mediante estas comparaciones de secuencias genéticas se obtienen conclusiones sobre:

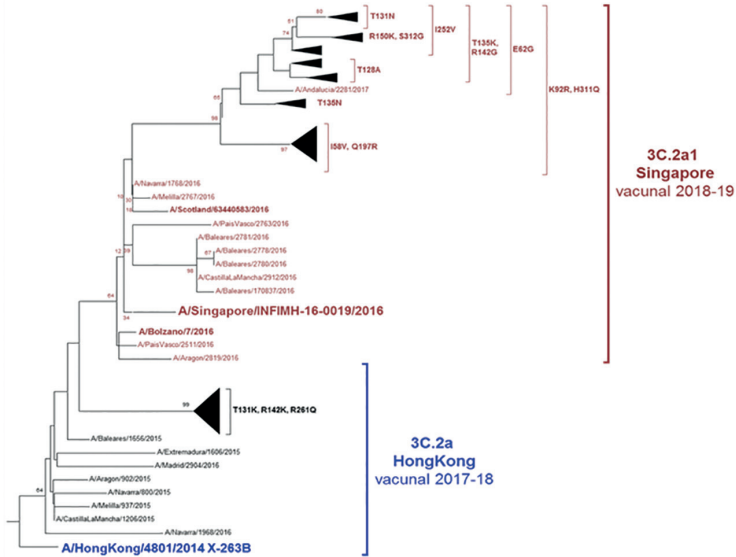
**Figura 3.4**

Grupos genéticos de los virus A/H3N2 circulantes en la temporada 2016-2017 y su relación antigénica con la cepa vacunal A/Hong Kong/4801/2014 variante X-263B producida en embrión de pollo.



**Figura 3.5**

Grupos genéticos definidos por aminoácidos localizados en lugares antígenicos de la hemaglutinina de los virus A/H3N2 circulantes en España en la temporada 2017-2018 y relación con la nueva cepa vacunal definida para la temporada 2018-2019, A/Singapore/INFIMH-16-0019/2016, dado que son antigénicamente diferentes del componente vacunal A/Hong Kong/4801/2014 variante X-263B utilizado en las vacunas trivalentes y tetravalentes.



- Cómo los virus se «relacionan» entre sí.
- Cómo los virus «evolucionan» a lo largo del tiempo.
- Cuáles son las variaciones genéticas o «mutaciones» que aparecen cuando los virus se diseminan ocasionando enfermedades más graves o desarrollando resistencia a los medicamentos antivirales.
- Qué grado de semejanza presenta un determinado virus de la gripe con el incluido como componente de la vacuna de la gripe.
- Cuáles son las adaptaciones que se producen en los virus animales que podrían causar un salto de la barrera interespecies e infectar al ser humano.

Las diferencias relativas que se observan tras el análisis de un grupo de secuencias procedentes de la caracterización genética de un grupo de virus de la gripe se organizan y se muestran gráficamente en un «árbol filogenético», que presenta la relación genética entre cada uno de los virus. Las secuencias se agrupan en función de la homología de los nucleótidos. Para la caracterización genética de los virus de la gripe circulantes en una temporada se realizan árboles filogenéticos que muestran la semejanza de los genes de HA o de los genes de NA de los virus. Cada secuencia de un virus ocupa un lugar en una rama del árbol, y el grado de diferencia genética (cantidad de diferencias en los nucleótidos) entre los virus se representa con la longitud de las ramas o líneas horizontales. Cuanto más alejados están los virus en el eje horizontal del árbol filogenético, más diferentes genéticamente son unos de otros.

La técnica de secuenciación denominada «método de Sanger» se utiliza tradicionalmente para conocer las características gené-

ticas y la evolución de los virus de la gripe como parte de la vigilancia virológica. La secuenciación de Sanger identifica la secuencia genética que predomina entre las variantes de los virus de la gripe, que ha sido seleccionada en un aislamiento viral. La aplicación de la secuenciación de Sanger directamente en muestras clínicas ha supuesto el estudio de los virus presentes en la muestra, aunque si en una población viral existen variantes en baja proporción, dichas variantes no pueden analizarse en la secuencia final. En muchos laboratorios del mundo se analiza un fragmento del segmento que codifica la HA y se obtienen secuencias parciales de ese segmento que son lo suficiente informativas como para realizar los análisis entre los virus y definir su homología o semejanza, y de este modo poder hacer la vigilancia de los virus que circulan en una determinada temporada (Figuras 3.3, 3.4 y 3.5). Desafortunadamente, mediante la secuenciación de Sanger es costoso obtener las secuencias del genoma completo de los virus grippales.

En los últimos años se han desarrollado y se están implementando los «sistemas de secuenciación de nueva generación» (NGS, por sus siglas en inglés), que amplían significativamente el volumen de secuencias obtenidas mediante una ultrasecuenciación o secuenciación profunda. A diferencia de la secuenciación de Sanger, los NGS utilizan fragmentos de ADN individualizados en una librería y segregados en un chip. Por lo tanto, los NGS revelan las variaciones genéticas que hay entre muchas partículas de virus diferentes en una sola muestra clínica. Estos métodos sirven para obtener datos sobre el genoma completo de los virus, entendiendo a estos como una población viral, y los datos deben ser analizados bioinformáticamente mediante plataformas integradoras<sup>11,12</sup>. Además, la información obtenida mejora el conocimiento de la deriva genética integral, ya que además de los genes que codifican HA y NA se secuencian los genes internos de manera completa.

El primer método de ultrasecuenciación que se comercializó fue el sistema 454 FLX (<http://www.454.com/>), desarrollado en 2005. En 2007 se comenzó a comercializar el sistema Illumina (<http://www.illumina.com>), basado en la secuenciación por síntesis y que ofrece plataformas versátiles (MySeq, HiSeq, MiniSeq y recientemente, en 2018, iSeq100). Existe una enorme expansión de nuevas plataformas, como Heliscope (<http://www.helicosbio.com/>) y Ion Torrent PGM (<http://www.iontorrent.com/>)<sup>13,14</sup>, o la tercera generación Pacific Biosciences (<http://www.pacificbiosciences.com/>). Muchos laboratorios han implementado estos sistemas, que han demostrado ser de enorme utilidad para la investigación de los virus de la gripe A/H1N1pdm de 2009, lo que supuso una revolución en el ámbito de la investigación genómica básica y aplicada<sup>4,15</sup>.

Actualmente, obtener el genoma completo de los virus gripales es asumible y se ha implementado de manera general en el estudio de su evolución, patogenia y ecología<sup>16-23</sup>.

### 3.4 Bases moleculares de la deriva antigénica

Los cambios o mutaciones que aparecen debido a la deriva genética o cambios menores (*drift* genético) favorecen la aparición de virus que escapan al control de los anticuerpos neutralizantes producidos en individuos infectados de manera natural o en individuos vacunados. El proceso que origina nuevas variantes del virus gripal se denomina «deriva antigénica».

Los estudios de antigenicidad se centran en las dos proteínas de superficie: HA y NA. Ambas han sido estudiadas intensamente desde el punto de vista genético o molecular por ser las proteínas más inmunógenas del virus. La HA es el mayor antígeno del virus de la gripe y el principal objetivo de la

respuesta inmunitaria. Es una proteína integral de membrana de tipo I compuesta por un dominio globular «cabeza» y un «tallo» más conservado. La cabeza globular es la región más variable y en ella se localiza una de las estructuras más importantes relacionadas con la patogenicidad de los virus gripales: el sitio de unión al receptor (RBS, *receptor binding site*)<sup>24</sup> (Figura 3.6). El RBS es una región conservada considerada el objetivo principal de los anticuerpos neutralizantes que tienen como fin prevenir la entrada del virus en la célula; de esta manera, su acción se desarrolla bloqueando la unión al receptor celular<sup>25</sup>. Por otra parte, la NA es una proteína integral de tipo II compuesta por cuatro subunidades idénticas. Cada subunidad tiene un dominio citoplasmático, un dominio transmembrana, un tallo y una cabeza globular<sup>26</sup>. A diferencia de la HA, el tallo es la región de la NA más variable<sup>27</sup>, y la cabeza, donde se encuentra el sitio de activación enzimática, es el más conservado. El estudio de los determinantes moleculares de la deriva antigénica ha sido prioritario durante décadas debido a que los cambios en estas proteínas, junto con la presión selectiva, permiten seleccionar las variantes antigénicas que escapan de la acción de los anticuerpos neutralizantes, y es por ello que el conocimiento de estos mecanismos resulta determinante para seleccionar los componentes de las vacunas antigripales y diseñar fármacos antivirales.

A pesar de que la gripe tipo A y la gripe tipo B tienen similitudes en cuanto a que son estacionales en el ser humano y causa de epidemias y pandemias, sus mecanismos de deriva antigénica son ligeramente diferentes.

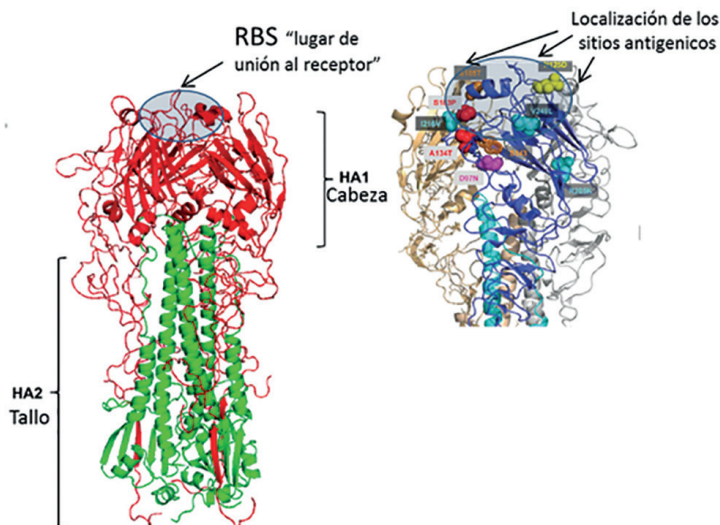
#### 3.4.1 Deriva antigénica en la gripe tipo A

El conocimiento de la deriva antigénica de la gripe tipo A es el más extenso, y gra-



Figura 3.6

Esquema de la morfología de la hemaglutinina, localización del lugar de unión al receptor y mutaciones que se acumulan a su alrededor o en lugares antigénicos. Ejemplo: virus A/H1N1pdm.



cias a las nuevas técnicas de secuenciación de nueva generación, el genoma completo de los virus ampliará en los próximos años diferentes aspectos básicos de la deriva antigénica. Uno de los trabajos clave para entender la deriva antigénica fue publicado en la década de 1980 y tenía como fin generar un mapa antigénico de la HA mediante un panel de anticuerpos neutralizantes<sup>28</sup>. En este trabajo se observó que los virus que presentaban determinadas mutaciones crecían en cultivos celulares, a diferencia de los que carecían de ellas y eran neutralizados. Mediante mutagénesis dirigida se pudo establecer un mapa de los lugares clave para generar virus que escapaban a la acción neutralizante de los anticuerpos. De esta manera se identificaron cinco lugares o «sitios antigénicos mayores», tanto para H1 como para H3, denominados Sa, Sb, Ca1, Ca2 y Cb para H1 y sitios A-E para H3. Las posiciones Sa, Sb y Ca2, y los sitios A y B de H3, se solapan con la región de unión al receptor<sup>29-31</sup> (**Figura 3.6**).

La aparición de sitios de N-glicosilación es la segunda estrategia que usa el virus de la gripe para evadir el sistema inmunitario, y al igual que los sitios antigénicos, suceden en el mismo lugar: las zonas próximas al RBS. La adición de glicanos en determinadas posiciones, incluidas algunas adyacentes al RBS<sup>32</sup>, puede enmascarar o modificar los sitios antigénicos y facilitar la evasión inmunitaria<sup>33,34</sup>. Esta estrategia de evasión es utilizada por el virus y se ha comprobado que el número de N-glicosilaciones se ha incrementado sustancialmente tanto en A/H1N1 como en A/H3N2 a lo largo del tiempo. Como ejemplo, el número de sitios de N-glicosilación por promotor aumentó de uno a tres desde 1918 hasta 1957<sup>34,35</sup> en el caso de virus A/H1N1, y entre 1968 y 2003 un total de 14 posiciones se han asociado con N-glicosilaciones en A/H3N2 circulantes<sup>32</sup>. Pese a su utilidad para el virus, esta estrategia debe ser equilibrada porque la hiperglicosilación de un sitio puede impedir una correcta unión y fusión del virus con la célula del hospedador, causándole una

pérdida total o parcial de su eficacia biológica<sup>36</sup>. A pesar de ser una estrategia de evasión, los cambios en la glicosilación en estas zonas no coinciden con los momentos de cambio de grupo antigénico<sup>34,37</sup>. Aunque es conocido que la N-glicosilación de sitios próximos al RBS puede introducir variaciones, como por ejemplo cambios en la virulencia<sup>34</sup>, y permitir la evasión inmunitaria, no se le considera aún causa directa de la deriva antigénica.

Durante el último medio siglo, las HA de los virus humanos A/H3N2 han acumulado al menos 75 sustituciones de aminoácidos. Los lugares donde se han producido cambios se encuentran nuevamente próximos al RBS, por lo que el conocimiento de los 131 aminoácidos que componen estas zonas próximas es de enorme importancia<sup>38-40</sup>. El estudio de los sitios antigénicos próximos al RBS mostró que un solo cambio en siete posiciones que se encontraban en los sitios A y B en la H3 (145, 155, 156, 158, 159, 189 y 193) ha sido determinante para que se produjesen cambios antigénicos de 1968 a 2003 en H3N2<sup>40</sup>. Y esto no solo ocurre en el virus H3N2 humano, sino que este fenómeno se replica tanto en los virus de la gripe aviar A/H5N1 como en los de la gripe equina A/H3N8 y porcina A/H3N2<sup>41-43</sup>.

Los anticuerpos cuyo objetivo es el tallo de la HA son menos prevalentes que los de la cabeza globular, ya que sus epítomos son más inaccesibles; por ello son necesarias infecciones múltiples para generar anticuerpos contra el tallo<sup>44</sup>. Actualmente, esta zona del tallo es muy interesante para el diseño de nuevas vacunas antigripales, ya que al ser más conservada la protección que generan los anticuerpos puede ofrecer protección frente a distintos subtipos de la gripe tipo A.

Como conclusión, en el RBS se localizan las posiciones clave para entender el mecanismo de deriva antigénica, pues los cambios de aminoácidos en estas siete posiciones crí-

ticas son la única causa bien caracterizada de la deriva antigénica, lo que indica que las posibilidades de cambio antigénico son más restringidas de lo que pensábamos.

Es interesante señalar que la deriva antigénica también se ha asociado a cambios en la NA viral, aunque esto ha sido menos estudiado debido a la inmunodominancia de la HA, que es más abundante en la superficie viral que la NA, aunque ambas son igualmente inmunógenas<sup>45</sup>. Los cambios que ocurren en la NA están asociados con los producidos en la HA. Ambas proteínas deben tener un equilibrio funcional para una eficiente replicación viral, y cuando se desequilibran por la aparición de cambios en alguna de ellas, el número de mutaciones en la otra aumenta para restaurar dicho equilibrio<sup>46</sup>. Debido a este fenómeno de mutaciones compensatorias, el estudio antigénico de la NA es complementario al de la HA. Es de enorme interés profundizar en el conocimiento sobre cómo este equilibrio HA-NA influye en la deriva antigénica global de los virus gripales y cómo puede proporcionar información relevante sobre este mecanismo.

### 3.4.2 Deriva antigénica en la gripe tipo B

La gripe tipo B ha sido menos estudiada que la gripe tipo A, posiblemente porque las epidemias que causa sean menos frecuentes, menos intensas, tardías en su aparición o quizás menos graves. La gripe tipo B no presenta subtipos, pero se han definido dos linajes cocirculantes genéticamente y antigénicamente distintos: el linaje Victoria y el linaje Yamagata<sup>47</sup>. Estos linajes tienen una compleja historia epidemiológica desde su divergencia, su cocirculación hasta 2002 y su alternancia.

El linaje Victoria está formado por un único grupo o clado, a diferencia del linaje



Yamagata, constituido por dos clados co-circulantes (clados 2 y 3) con importantes implicaciones en la dinámica de estos virus. Como ejemplo, en Malasia, en 2013, el linaje circulante fue definido como clado 2 tras haber circulado la temporada anterior el linaje Victoria, y en 2014 el linaje Yamagata continuó circulando, pero fue el clado 3 el dominante<sup>48</sup>. La existencia de los dos linajes diferentes genéticamente y antigénicamente cocirculando, y de virus pertenecientes a distintos grupos genéticos, puede complicar mucho la selección del componente vacunal para la siguiente temporada gripal. Como en la gripe tipo A, esta transición de un cluster antigénico a otro se debe a una única sustitución de aminoácido en una región cercana al RBS<sup>40</sup>. Aunque la causa de la deriva antigénica en ambos virus es la misma, hay una diferencia principal entre los dos tipos de virus, y es que el tipo B tiene tasas de mutación y de sustitución de aminoácidos más lentas, pero más diversas, que el tipo A<sup>49</sup>. Existen diferentes hipótesis que intentan explicar las causas de que la tasa de mutación sea más lenta. Una de ellas propone que la tasa de error de la RdRp es menor que la de los virus de la gripe A<sup>50</sup>, y otra propone que son menos inmunógenos que los de la gripe A<sup>51</sup> y que su avidez por el ácido siálico es menor que en la gripe A<sup>49,52</sup>. Estas hipótesis, junto con que las epidemias por estos virus son menos frecuentes, pueden explicar que la deriva antigénica de la gripe B sea más lenta.

Hay que destacar que los virus tipo B presentan una alta tasa de reorganización y es frecuente la reorganización interlinajes e intralíneas. Además, las inserciones y las deleciones de nucleótidos pueden ser la causa de su diversificación genética y antigénica<sup>53</sup>.

Es necesario ampliar y completar el conocimiento de los procesos de deriva antigénica y las dinámicas evolutivas en la gripe tipo B,

ya que su capacidad de diversificarse podría complicar la selección de los componentes vacunales, tanto en la vacuna trivalente como en la tetravalente.

### 3.5 La caracterización genética y antigénica de los virus de la gripe

Los datos genéticos se utilizan junto con los datos de la caracterización antigénica de los virus para determinar qué virus de la vacuna deben elegirse para las próximas vacunas de la gripe que se distribuirán tanto en el hemisferio norte y como en el hemisferio sur. A través del sistema de vigilancia se obtienen los virus circulantes, y sus secuencias se comparan con las de los genes HA y NA de los virus incluidos en la vacuna de una determinada temporada. Así se evalúan las diferencias que pueden aparecer por la deriva genética de los virus gripales. No es raro demostrar que, durante el transcurso de una o varias temporadas epidémicas, los virus que circulan presentan cambios genéticos que les hacen ser diferentes de los virus incluidos en la vacuna correspondiente. Si esto sucede es necesario seleccionar otro virus como componente de la vacuna de la próxima temporada gripal (Tablas 3.1 y 3.2). La caracterización antigénica se refiere al análisis de la reacción de un virus ante la presencia de anticuerpos para ayudar a evaluar cuánto está relacionado con otro virus.

Actualmente, la estimación de la distancia antigénica basada en la secuencia de una determinada proteína es un sistema de enorme utilidad para la investigación de los virus gripales, y como las bases de datos contienen cientos de miles de secuencias de virus gripales, la predicción de la discordancia (mismatch) antigénica (Tablas 3.1 y 3.2) da lugar a estudios que modelizan estos aspectos y los aplican a la predicción de la efectividad vacunal y de cómo los anticuerpos producidos

**Tabla 3.1**

**Circulación AH1N1pdm09**  
**Relación genética y antigénica**

Año	Subgrupos genéticos	Cepa vacunal
2009-10	inicial	A/California/07/09
2010-11	5, 6, 7	
2011-12	5, 6, 7	
2012-13	6A	
2013-14	6B	
2014-15	6B	
2015-16	6B.1 y 6B.2	
2016-17	6B.1	
2017-18	6B.1	A/Michigan/45/2015*
2018-19	??	

\*Sustitución del componente vacunal: baja reactividad de los sueros humanos tras la vacunación

**Tabla 3.2**

**Circulación AH3N2**  
**Diversidad genética y Deriva antigénica**

Año	Subgrupos	Cepa vacunal
2010-11	1	A/Perth/16/2009
2011-12	3A,3B,3C	
2012-13	3A,3B,3C.1, 3C.2	
2013-14	3C.2, 3C.3	A/Victoria/361/2011 (3C.1)
2014-15	3C.2a, 3C.3a	A/Texas/50/2012 (3C.1)
2015-16	3C.2, 3C.2a, 3C.3, 3C.3a, después Feb 3C.3b	A/Switzerland/9715293/2013 (3C.3a)
2016-17	3C.2a,3C.2a1, 3C.3a	A/HongKong/4801/2014 (3C.2a)
2017-18	3C.2a1	
2018-19	??	A/Singapore/INFIMH-16-0019/2016 (3C.2a1)

ante una primoinfección gripal condicionan la respuesta del individuo por las reacciones cruzadas de sus anticuerpos<sup>54</sup>.

### 3.5.1 Ejemplo: el virus de la gripe A/H3N2 (Figuras 3.4 y 3.5)

Los virus de la gripe A/H3N2 derivaron en 2015<sup>55</sup>, y durante la temporada 2015-2016 se observó en el hemisferio norte una diversificación genética de la HA en las cepas circulantes; en concreto, las mutaciones se encontraban en la región HA1 del gen de la HA<sup>56,57</sup>. En el comienzo de la temporada se reconocían dos grupos genéticos: uno de ellos integrado por los virus similares al virus vacunal A/HongKong/4801/2014 (grupo 3C.2a) y otro constituido por virus con la mutación N171K (grupo 3C.2a1), representado por el virus A/Bolzano/7/2016. Ambos grupos no presentaban diferencias antigénicas entre ellos. A medida que fue transcurriendo la temporada, la incorporación de nuevas mutaciones permitió el reconocimiento de nuevos subgrupos o clados genéticos. En el grupo 3C.2a se han identificado dos subgrupos: uno de ellos caracterizado por la presencia de las mutaciones N121K y S144K (a menudo también N122D y S262N) y el otro caracterizado por las mutaciones T131K y R142K (a veces también R261Q). En el grupo 3C.2a1, la diversidad fue aún mayor y se identificaron hasta tres subgrupos diferentes atendiendo a la presencia de las mutaciones T135K, I140M y R142G. La mayoría de estas sustituciones de aminoácidos se localizan en el sitio antigénico A de la HA, y además, en el caso de N122D y T135K, llevan implícita la pérdida de un potencial lugar de glicosilación.

La extraordinaria diversidad de los virus A/H3N2 circulantes probablemente sea el distintivo más destacado en las últimas temporadas gripales en las que este virus ha sido el predominante. Este hecho supone un

auténtico desafío para la selección del virus vacunal, situación agravada por la circunstancia de que los cambios introducidos en la HA de estos virus durante los últimos años han afectado al sitio de reconocimiento del receptor celular, mermando la capacidad de hemaglutinación de estos virus y, por tanto, limitando los estudios antigénicos que tradicionalmente se realizan mediante inhibición de la hemaglutinación<sup>58</sup>.

## 3.6 Cambios para la evolución viral y presión selectiva

A pesar de que asociamos la gripe con un virus que cambia rápidamente, la realidad es que su evolución es lenta y la formación de nuevas variantes del virus es difícil. Sorprendentemente, la tasa de aparición de nuevos grupos antigénicos calculada mediante modelos matemáticos es cada 3,3 años<sup>59</sup>. Como ya hemos mencionado, solo siete posiciones en la HA son las causantes de cambios antigénicos, lo que reduce mucho las posibilidades de generar nuevas variantes antigénicas del virus. Que estas posiciones se encuentren alrededor de la zona del RBS implica que cualquier cambio que se produzca debe evitar alterar la función de la proteína, y por tanto la probabilidad de que ocurran cambios se reduce aún más. Los estudios de modelos dinámicos de epidemiología y evolución de la gripe<sup>60</sup> muestran que la ratio de cambios antigénicos puede ser limitada por la necesidad de adquirir muchas mutaciones para que, aunque se altere su antigenicidad, no se modifique su eficacia biológica. Incluso en ausencia de las limitaciones anteriores, la oportunidad de que se seleccione una variante es baja porque es necesario alcanzar una densidad suficiente de virus variante para que la selección opere. Si unimos la naturaleza aguda de la infección con la reducida posibilidad de que se produzca una variante y la acción de control que ejerce el sistema inmunitario, se observa que la producción de nuevas variantes es un fenómeno difícil y

más raro de lo que se podría creer en un principio<sup>61</sup>. Si las condiciones anteriores limitan la producción de nuevas variantes, la transmisión de estas nuevas variantes supone una nueva barrera para su establecimiento como virus circulante. Las nuevas variantes deben ser capaces de replicarse y estar en una cantidad suficiente que permita la transmisión; es conocido que enfrentarse a las barreras de la inmunidad innata supone una pérdida de la diversidad generada<sup>62</sup>. Además del proceso de transmisión, la naturaleza estacional de la infección dificulta que estas nuevas variantes compitan con las existentes para ser el germen de la siguiente epidemia.

A pesar de todo esto, el conocimiento evolutivo sobre la gripe aumentará gracias a las técnicas de secuenciación de nueva generación que permitirán conocer las mutaciones durante la infección en la célula del hospedador cuando esta nueva variante se transmita a otros individuos, permitiendo predecir patrones de evolución viral que nos darán una información más completa para mejorar el control del virus.

## Bibliografía

- Chen W, Calvo PA, Malide D, Gibbs J, Schubert U, Bacik I, et al. A novel influenza A virus mitochondrial protein that induces cell death. *Nat Med*. 2001;7:1306-12.
- Hayashi T, MacDonald LA, Takimoto T. Influenza A virus protein PA-X contributes to viral growth and suppression of the host antiviral and immune responses. *J Virol*. 2015;89:6442-52.
- Smith F.I., Palese P. (1989) Variation in Influenza Virus Genes. In: Krug R.M. (eds) *The Influenza Viruses. The Viruses*. Springer, Boston, MA. ~~Mal citada, incompleta~~
- Nelson M, Spiro D, Wentworth D, Beck E, Fan J, Ghedin E, et al. The early diversification of influenza A/H1N1pdm. *PLoS Curr*. 2009;1:RRN1126.
- Hill NJ, Ma EJ, Meixell BW, Lindberg MS, Boyce WM, Runstadler JA. Transmission of influenza reflects seasonality of wild birds across the annual cycle. *Ecol Lett*. 2016;19:915-25.
- Venkatesh D, Poen MJ, Bestebroer TM, Scheuer RD, Vuong O, Chkhaidze M, et al. Avian influenza viruses in wild birds: virus evolution in a multihost ecosystem. *J Virol*. 2018;92:e00433-18.
- Labadie T, Batéjat C, Manuguerra J-C, Leclercq I. Influenza virus segment composition influences viral stability in the environment. *Front Microbiol*. 2018;9:1496.
- Marshall N, Priyamvada L, Ende Z, Steel J, Lowen AC. Influenza virus reassortment occurs with high frequency in the absence of segment mismatch. *PLoS Pathog*. 2013;9(6):e1003421.
- Garten RJ, Davis CT, Russell CA, Shu B, Lindstrom S, Balish A, et al. Antigenic and genetic characteristics of the early isolates of swine-origin 2009 A(H1N1) influenza viruses circulating in humans. *Science*. 2009;325:197-201.
- Shinde V, Bridges CB, Uyeki TM, Shu B, Balish A, Xu X, et al. Triple-reassortant swine influenza A (H1) in humans in the United States, 2005-2009. *N Engl J Med*. 2009;360:2616-25.
- Shepard SS, Meno S, Bahl J, Wilson MM, Barnes J, Neuhaus E. Viral deep sequencing needs an adaptive approach: IRMA, the iterative refinement meta-assembler. *BMC Genomics*. 2016;17:708.

12. Borges V, Pinheiro M, Pechirra P, Guio-  
mar R, Gomes JP. INSaFLU: an auto-  
mated open web-based bioinformatics  
suite “from-reads” for influenza who-  
le-genome-sequencing-based surveillan-  
ce. *Genome Med.* 2018;10:46.
13. Quail MA, Smith M, Coupland P, Otto  
TD, Harris SR, Connor TR, et al. A tale  
of three next generation sequencing plat-  
forms: comparison of Ion Torrent, Pacific  
Biosciences and Illumina MiSeq sequen-  
cers. *BMC Genomics.* 2012;13:341.
14. Jünemann S, Sedlazeck FJ, Prior K, Al-  
bersmeier A, John U, Kalinowski J, et  
al. Updating benchtop sequencing per-  
formance comparison. *Nat Biotechnol.*  
2013;31:294-6.
15. Greninger AL, Chen EC, Sittler T,  
Scheinerman A, Roubinian N, Yu G,  
et al. A metagenomic analysis of pande-  
mic influenza A (2009 H1N1) infection  
in patients from North America. *PLoS  
One.* 2010;5:e13381.
16. Meinel DM, Heinzinger S, Eberle U,  
Ackermann N, Schönberger K, Sing A.  
Whole genome sequencing identifies in-  
fluenza A H3N2 transmission and offers  
superior resolution to classical typing  
methods. *Infection.* 2018;46:69-76.
17. Yildiz S, Mazel-Sanchez B, Kandasamy  
M, Manicassamy B, Schmolke M. In-  
fluenza A virus infection impacts sys-  
temic microbiota dynamics and causes  
quantitative enteric dysbiosis. *Microbio-  
me.* 2018;6:9.
18. Qi W, Jia W, Liu D, Li J, Bi Y, Xie S, et  
al. Emergence and adaptation of a novel  
highly pathogenic H7N9 influenza virus  
in birds and humans from a 2013 hu-  
man-infecting low-pathogenic ancestor.  
*J Virol.* 2018;92(2).
19. Pichon M, Gaymard A, Josset L, Valette  
M, Millat G, Lina B, et al. Characteri-  
zation of oseltamivir-resistant influenza  
virus populations in immunosuppres-  
sed patients using digital-droplet PCR:  
comparison with qPCR and next ge-  
neration sequencing analysis. *Antiviral  
Res.* 2017;145:160-7.
20. Sobel Leonard A, McClain MT, Smith  
GJ, Wentworth DE, Halpin RA, Lin  
X, et al. Deep sequencing of influenza  
A virus from a human challenge study  
reveals a selective bottleneck and only li-  
mited intrahost genetic diversification. *J  
Virol.* 2016;90:11247-58.
21. Zhao J, Liu J, Vemula SV, Lin C, Tan  
J, Ragupathy V, et al. Sensitive detec-  
tion and simultaneous discrimination  
of influenza A and B viruses in naso-  
pharyngeal swabs in a single assay using  
next-generation sequencing-based diag-  
nostics. *PLoS One.* 2016;11:e0163175.
22. Seong M-W, Cho SI, Park H, Seo SH,  
Lee SJ, Kim EC, et al. Genotyping in-  
fluenza virus by next-generation deep se-  
quencing in clinical specimens. *Ann Lab  
Med.* 2016;36:255-8.
23. Mei K, Liu G, Chen Z, Gao Z, Zhao  
L, Jin T, et al. Deep sequencing reveals  
the viral adaptation process of environ-  
ment-derived H10N8 in mice. *Infect  
Genet Evol.* 2016;37:8-13.
24. Ndifon W, Wingreen NS, Levin SA.  
Differential neutralization efficiency of  
hemagglutinin epitopes, antibody in-  
terference, and the design of influen-  
za vaccines. *Proc Natl Acad Sci U S A.*  
2009;106:8701-6.
25. Barbey-Martin C, Gigant B, Bizebard  
T, Calder LJ, Wharton SA, Skehel JJ, et  
al. An antibody that prevents the hema-

- gglutinin low pH fusogenic transition. *Virology*. 2002;294:70-4.
26. Air GM. Influenza neuraminidase. *Influenza Other Respir Viruses*. 2012;6:245-56.
  27. Wu C-Y, Lin CW, Tsai TI, Lee CD, Chuang HY, Chen JB, et al. Influenza A surface glycosylation and vaccine design. *Proc Natl Acad Sci U S A*. 2017;114:280-5.
  28. Caton AJ, Brownlee GG, Yewdell JW, Gerhard W. The antigenic structure of the influenza virus A/PR/8/34 hemagglutinin (H1 subtype). *Cell*. 1982;31:417-27.
  29. Wiley DC, Wilson IA, Skehel JJ. Structural identification of the antibody-binding sites of Hong Kong influenza haemagglutinin and their involvement in antigenic variation. *Nature*. 1981;289:373-8.
  30. Skehel JJ, Stevens DJ, Daniels RS, Douglas AR, Knossow M, Wilson IA, et al. A carbohydrate side chain on hemagglutinins of Hong Kong influenza viruses inhibits recognition by a monoclonal antibody. *Proc Natl Acad Sci U S A*. 1984;81:1779-83.
  31. Gerhard W, Yewdell J, Frankel ME, Webster R. Antigenic structure of influenza virus haemagglutinin defined by hybridoma antibodies. *Nature*. 1981;290:713-7.
  32. Blackburne BP, Hay AJ, Goldstein RA. Changing selective pressure during antigenic changes in human influenza H3. *PLOS Pathog*. 2008;4:e1000058.
  33. Schulze IT. Effects of glycosylation on the properties and functions of influenza virus hemagglutinin. *J Infect Dis*. 1997;176(Suppl 1):S24-8.
  34. Medina RA, Stertz S, Manicassamy B, Zimmermann P, Sun X, Albrecht RA, et al. Glycosylations in the globular head of the hemagglutinin protein modulate the virulence and antigenic properties of the H1N1 influenza viruses. *Sci Transl Med*. 2013;5:187ra70.
  35. Zhang M, Gaschen B, Blay W, Foley B, Haigwood N, Kuiken C, et al. Tracking global patterns of N-linked glycosylation site variation in highly variable viral glycoproteins: HIV, SIV, and HCV envelopes and influenza hemagglutinin. *Glycobiology*. 2004;14:1229-46.
  36. Tsuchiya E, Sugawara K, Hongo S, Matsuzaki Y, Muraki Y, Li ZN, et al. Effect of addition of new oligosaccharide chains to the globular head of influenza A/H2N2 virus haemagglutinin on the intracellular transport and biological activities of the molecule. *J Gen Virol*. 2002;83:1137-46.
  37. Tate MD, Job ER, Brooks AG, Reading PC. Glycosylation of the hemagglutinin modulates the sensitivity of H3N2 influenza viruses to innate proteins in airway secretions and virulence in mice. *Virology*. 2011;413:84-92.
  38. Wang X, Ilyushina NA, Lugovtsev VY, Bovin NV, Couzens LK, Gao J, et al. Amino acids in hemagglutinin antigenic site B determine antigenic and receptor binding differences between A(H3N2)v and ancestral seasonal H3N2 influenza viruses. *J Virol*. 2017;91(2).
  39. Popova L, Smith K, West AH, Wilson PC, James JA, Thompson LF, et al. Immunodominance of antigenic site B over site A of hemagglutinin of recent H3N2 influenza viruses. *PLoS One*. 2012;7:e41895.



40. Koel BF, Burke DF, Bestebroer TM, van der Vliet S, Zondag GC, Vervaeke G, et al. Substitutions near the receptor binding site determine major antigenic change during influenza virus evolution. *Science*. 2013;342:976-9.
41. Koel BF, van der Vliet S, Burke DF, Bestebroer TM, Bharoto EE, Yasa IW, et al. Antigenic variation of clade 2.1 H5N1 virus is determined by a few amino acid substitutions immediately adjacent to the receptor binding site. *MBio*. 2014;5:e01070-14.
42. Abente EJ, Santos J, Lewis NS, Gauger PC, Stratton J, Skepner E, et al. The molecular determinants of antibody recognition and antigenic drift in the H3 hemagglutinin of swine influenza A virus. *J Virol*. 2016;90:8266-80.
43. Lewis NS, Daly JM, Russell CA, Horton DL, Skepner E, Bryant NA, et al. Antigenic and genetic evolution of equine influenza A (H3N8) virus from 1968 to 2007. *J Virol*. 2011;85:12742-9.
44. Nachbagauer R, Choi A, Izikson R, Cox MM, Palese P, Krammer F. Age dependence and isotype specificity of influenza virus hemagglutinin stalk-reactive antibodies in humans. *MBio*. 2016;7:e01996-15.
45. Akram A, Inman RD. Immunodominance: a pivotal principle in host response to viral infections. *Clin Immunol*. 2012;143:99-115.
46. Mitnaul LJ, Matrosovich MN, Castrucci MR, Tuzikov AB, Bovin NV, Kobasa D, et al. Balanced hemagglutinin and neuraminidase activities are critical for efficient replication of influenza A virus. *J Virol*. 2000;74:6015-20.
47. Rota PA, Wallis TR, Harmon MW, Rota JS, Kendal AP, Nerome K. Cocirculation of two distinct evolutionary lineages of influenza type B virus since 1983. *Virology*. 1990;175:59-68.
48. Oong XY, Ng KT, Lam TT, Pang YK, Chan KG, Hanafi NS, et al. Epidemiological and evolutionary dynamics of influenza B viruses in Malaysia, 2012-2014. *PLoS One*. 2015;10:e0136254.
49. Bedford T, Riley S, Barr IG, Broor S, Chadha M, Cox NJ, et al. Global circulation patterns of seasonal influenza viruses vary with antigenic drift. *Nature*. 2015;523:217-20.
50. Nobusawa E, Sato K. Comparison of the mutation rates of human influenza A and B viruses. *J Virol*. 2006;80:3675-8.
51. Monto AS, Maassab HF. Ether treatment of type B influenza virus antigen for the hemagglutination inhibition test. *J Clin Microbiol*. 1981;13:54-7.
52. Hensley SE, Das SR, Bailey AL, Schmidt LM, Hickman HD, Jayaraman A, et al. Hemagglutinin receptor binding avidity drives influenza A virus antigenic drift. *Science*. 2009;326:734-6.
53. McCullers JA, Wang GC, He S, Webster RG. Reassortment and insertion-deletion are strategies for the evolution of influenza B viruses in nature. *J Virol*. 1999;73:7343-8.
54. Anderson CS, McCall PR, Stern HA, Yang H, Topham DJ. Antigenic cartography of H1N1 influenza viruses using sequence-based antigenic distance calculation. *BMC Bioinformatics*. 2018;19:51.
55. Broberg E, Snacken R, Adlhoch C, Beauté J, Galinska M, Pereyaslov D, et al. Start

- of the 2014/15 influenza season in Europe: drifted influenza A(H3N2) viruses circulate as dominant subtype. *Euro Surveill.* 2015;20(4).
56. Harvala H, Frampton D, Grant P, Raffle J, Ferns RB, Kozlakidis Z, et al. Emergence of a novel subclade of influenza A(H3N2) virus in London, December 2016 to January 2017. *Euro Surveill.* 2017;22(8).
  57. Kissling E, Rondy M; I-MOVE/I-MOVE+ study team. Early 2016/17 vaccine effectiveness estimates against influenza A(H3N2): I-MOVE multicentre case control studies at primary care and hospital levels in Europe. *Euro Surveill.* 2017;22(7).
  58. Lin Y, Wharton SA, Whittaker L, Dai M, Ermetal B, Lo J, et al. The characteristics and antigenic properties of recently emerged subclade 3C.3a and 3C.2a human influenza A(H3N2) viruses passaged in MDCK cells. *Influenza Other Respir Viruses.* 2017;11:263-74.
  59. Smith DJ, Lapedes AS, de Jong JC, Bestebroer TM, Rimmelzwaan GF, Osterhaus AD, et al. Mapping the antigenic and genetic evolution of influenza virus. *Science.* 2004;305:371-6.
  60. Koelle K, Cobey S, Grenfell B, Pascual M. Epochal evolution shapes the phylodynamics of interpandemic influenza A (H3N2) in humans. *Science.* 2006;314:1898-903.
  61. McCrone JT, Woods RJ, Martin ET, Malosh RE, Monto AS, Luring AS. The evolutionary dynamics of influenza A virus within and between human hosts. *bioRxiv* 176362 (2017). Disponible en: <https://www.biorxiv.org/content/early/2017/11/25/176362>
  62. Poon LL, Song T, Rosenfeld R, Lin X, Rogers MB, Zhou B, et al. Quantifying influenza virus diversity and transmission in humans. *Nat Genet.* 2016;48:195-200.